DOCUMENT RESUME

ED 165 114 CS 004 610

| | |
|---|---|
| AUTHOR | Reder, Lynne M. |
| TITLE | Comprehension and Retention of Prose: A Literature Review. Technical Report No. 108. |
| INSTITUTION | Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading. |
| SPONS AGENCY | National Inst. of Education (DHEW), Washington, D.C. |
| PUB DATE | Nov 78 |
| CONTRACT | 400-76-0116 |
| NOTE | 116p. |
| EDRS PRICE | MF-$0.83 HC-$6.01 Plus Postage. |
| DESCRIPTORS | *Discourse Analysis; *Literature Reviews; Memory; Models; *Prose; *Reading Comprehension; *Reading Research; Recall (Psychological); Retention Studies; Semantics; Structural Analysis; Structural Grammar |
| IDENTIFIERS | *Center for the Study of Reading (Illinois); *Schemata |

ABSTRACT

This literature review explores research that has been done on prose comprehension and retention. The first section discusses investigations of factors that affect the amount of recall, including instructionally relevant experiments. research using additional context as external aids, and investigations that test the generalizability of verbal learning results to connected discourse. The second section discusses representation of text structures, including representations of logical relations, representations concerned with more complex relations, formal models of text structure, and story grammars. A third section discusses the use of world knowledge to aid comprehension, including schemata, frames, and prototypes; necessary inferences; scripts, plans, and goals; the script-elaboration model; and experimental evidence in support of the elaboration theory. The paper closes with general conclusions and implications for education. (TJ)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED165114

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Technical Report No. 108

# COMPREHENSION AND RETENTION OF PROSE: A LITERATURE REVIEW

Lynne M. Reder

Carnegie-Mellon University
Pittsburgh, Pennsylvania

November 1978

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02138

The research reported herein was supported by the National Institute of Education under Contract No. US-NIE-C-400-76-0116.

CS006610

TABLE OF CONTENTS

Introduction..... 1

1. Investigations of Factors that Affect Amount of Recall..... 5
   1.1 Focus on Instructionally Relevant Experiments..... 5
   1.2 Research Using Additional Context as External Aids..... 11
   1.3 Investigations that test the Generalizability of Verbal Learning Results to Connected Discourse..... 16
   1.4 Discussion..... 20

2. Representations of Text Structures..... 21
   2.1 Representations of Logical Relations..... 22
      2.1.1 Dawes..... 22
      2.1.2 Crothers..... 23
      2.1.3 Frederiksen..... 25
   2.2 Representations Concerned with More Complex Relations..... 28
      2.2.1 Frederiksen..... 28
      2.2.2 Kintsch..... 30
      2.2.3 Meyer..... 35
   2.3 Formal Models of Text Structure..... 40
      2.3.1 Theory of Discourse and Macro-structures..... 41
   2.4 Story Grammars..... 46
      2.4.1 Rumelhart..... 46
      2.4.2 Thorndyke..... 50
      2.4.3 Mandler and Johnson..... 51
      2.4.4 Stein and Glenn..... 53
      2.4.5 Discussion of Story Grammars..... 55

3. Use of World Knowledge to Aid Comprehension..... 58
   3.1 Schemata, Frames and Prototypes..... 58
   3.2 Necessary Inferences..... 60
      3.2.1 Clark..... 61
      3.2.2 Schank's Model for Drawing Inferences..... 63
   3.3 Scripts, Plans and Goals..... 65
   3.4 The Script-Elaboration Model..... 70
   3.5 Experimental Evidence in Support of the Elaboration Theory..... 74
   3.6 Discussion..... 80

4. General Conclusions and Implications for Education..... 82
   4.1 Summing Up..... 84

Reference Note..... 86

Bibliography..... 87

Footnotes.....102

## INTRODUCTION

In the last 5 years, an interest in studying prose comprehension and retention has been rekindled among cognitive psychologists. The psychologist generally credited with first studying memory for stories was Bartlett (1932). His early research is still frequently cited and still influences the field. In his experiments, subjects read stories from different cultures, which they later recalled. The main finding was that recall was inaccurate; elements of the story which were poorly understood for cultural reasons were distorted in recall to conform better with his subjects' knowledge. He considered his results incompatible with the notion of "fixed lifeless traces" and instead posited that remembering is a constructive process based on general impressions and attitudes.

For approximately 30 years, since Bartlett's seminal work, psychologists left the investigation of reading to researchers in education.[1] Instead, psychologists looked at memory for lists of nonsense syllables, words and unrelated sentences. One reason psychologists may have shied away from studying prose is the problems inherent in using large units of analysis. With large units of analysis, experimenters often feel that they have less control over the subject's processing of the material. Thorndyke (1975), for example, states that language comprehension involves "integration of incoming language into the situation or context..., use of previously learned world knowledge, the generation of inferences...(and) of expectations...(p.2)." However, this feeling of more control with CVC's is illusory, I think: To the extent that subjects are successful at remembering CVC's, word lists, etc., they are using their long-term memory in unpredictable

ways to create mnemonics and integrate them into a "meaningful structure," (Montague, Adams & Kiess, 1966; Prytulak, 1971). If anything, the experimenter should have a better idea of the nature of subjects' processing with paragraph material.

Another reason experimenters have shied away from using prose material is the difficulty in determining the similarity and differences among passages. This makes comparisons across experiments that use prose difficult. There have been two main dependent measures employed in the study of prose comprehension and retention: question answering and free recall of the passage. Anderson (1972) has argued that not enough attention has been placed on the selection of questions that are intended to measure comprehension. Just as it is difficult to evaluate the similarity of prose passages, it is difficult to score recall of prose in a meaningful way to determine how much has been retained.

The new willingness to study prose comprehension, despite its problems of imprecision is due to several factors. First of all, some investigators have lost faith in the idea that results from simple laboratory experiments will generalize to natural materials.[2] Second, it is now easier to conduct research and develop theories related to reading than it was a few years ago, due to the developments in computer science and linguistics.

Computers are an aid to the study of prose in two ways. Paragraphs can now be presented to the subject "on-line" (on the computer TV monitors), each sentence displayed for only the time a specific subject needs to read it. Reading times can be recorded, questions can be asked at any desired point, eye movements can be monitored, speed/accuracy in answering the questions can be manipulated and so on. Computers are also an aid to

theoretical development as they allow simulation of prose processing. It would be very difficult to develop a model of prose processing in much detail without implementing it on a computer as the complexity is not conceptually tractable. Another reason theorists have returned to prose processing is the concomitant development of text grammars in linguistics (e.g., Lakoff, 1972). Psychologists and simulators use theoretical notions from linguistics to suggest experiments or ways to simulate. Linguistics is also influenced by work done in the other two fields.

There are other groups of investigators apart from cognitive psychologists who have been studying prose. Researchers in the field of education have been exploring this issue for many years (e.g., Ausubel, Frase, Gagne, Rothkopf). Memory theorists who have uncovered a number of principles using the "verbal learning materials" (e.g., lists of nonsense syllables or words) have investigated to what extent the principles generalize to connected discourse.

One of the problems, of course, in studying prose comprehension is the need to pare down the morass of complexities associated with prose. Thus far, the strategy to deal with the enormous complexity seems to be to partition the problem of prose understanding into subproblems and only focus on one or two aspects, e.g., the structure and grammars of stories, how the "content" (words and topics) affects recall, how orienting tasks affect comprehension, or the role of knowledge bundles--scripts or frames--in understanding.

Hopefully, this problem can be partitioned into different aspects, but unless investigators are aware of all aspects of the problem, little success

is likely. For example, those concerned with developing story grammars should consider what kind of process model or mechanism would enable the reader to build a representation from a given story using the grammar. What world knowledge is needed to comprehend the story and how does the grammar use it? Those concerned with developing "comprehension" or process models necessarily assume an end representation for the story. Use of a specified story grammar would facilitate decisions concerning importance of various aspects of a story when parsed and influence expectations about material not yet processed from the text.

The first section of this paper will review some of the research that has focused on applied, practical questions of interest to education. The results from some of this research indicate the need for a more sophisticated understanding of text structures and internal representations of text. A review of current work on text grammars is given, followed by a discussion of work on processing models. Finally in the last section, we can see how some of the theoretical work has implications for educational problems/ questions, with which we began.[3]

# 1. INVESTIGATIONS OF FACTORS THAT AFFECT AMOUNT OF RECALL

## 1.1 Focus on Instructionally Relevant Experiments

An important application of research dealing with prose processing is how to improve people's ability to read. Various techniques have been tried to improve the student's comprehension and retention of reading material. Carroll (1971) provides an extensive review. Ausubel (1960) introduced the notion of "advance organizers." He thought that adjunct aids which gave the reader a preview of the content of the passage would improve the reader's organization of the material, thus leading to better comprehension and retention. The advance organizers should activate or develop appropriate cognitive structures to which subjects could anchor the incoming ideas. An example of his task is to give subjects a passage before the critical passage. The passage either contains "subsuming" concepts of the critical passage or is essentially irrelevant to it. He predicted that prior exposure to subsuming concepts of "anchors" would facilitate retention. Ausubel and his associates (e.g., Ausubel & Fitzgerald, 1962) have found data to be consistent with these notions to some extent; they admit, however, that an advance organizer's usefulness interacts with many other variables such as previous knowledge and verbal ability.

Although many researchers agree that Ausubel's work provided interesting speculations that stimulated others to pursue similar questions, he has been criticized for lack of appropriate experimental controls and lack of objective measures of his stimulus variables (see Frase, 1975). It is difficult to test the notion that previous exposure to the high level concepts of a passage will improve retention of lower level concepts when

no procedure for evaluating the structure of a passage has been proposed. Gagne and Wiegand (1970) are examples of investigators who put Ausubel's conclusions into doubt. They found that improvement in retention due to Ausubel's advance organizers may have had their facilitation in retrieval, not in storage. Recall was improved even when the related topic sentence was not given until just before the test. This is inconsistent with Ausubel's argument that improvement results from the ability to embed the information in pre-existing structures.[4]

The investigation of the use of adjunct aids in the comprehension of prose has been continued by a number of investigators. One such aid that has been explored thoroughly is the effect of questions on subsequent test performance (e.g., Anderson & Biddle, 1975; Frase, 1967, 1968, 1971, 1972, 1975; McGaw & Grotelueschen, 1972; Rothkopf, 1966, 1972; Rothkopf & Bisbicos, 1967; Watts & Anderson, 1971; see Anderson & Biddle for an extensive review). The typical dependent measure is how well a subject can answer a set of questions about a passage (call these test questions) as a function of having been asked questions about the passage earlier (call these priming questions). Examples of the independent variables are whether the priming questions are asked before the subject reads the passage, during the passage or after the passage; whether the priming questions are general or specific; whether the subject is given feedback on the priming questions.

The majority of the studies in the field have found that subjects do better at answering a test question if they have been asked that same question earlier. The improvement obtains regardless of whether the priming question was given before the passage was read or afterwards and whether or not the subject was given feedback about his answer. Performance is best if

the question was asked before the passage was read. In the case where priming questions ask about information different from that to be tested, the findings are mixed: subjects do better on the to-be-tested questions than controls who have not seen priming questions only when the priming questions were asked after the passage was read. If the priming questions were asked prior to reading the text, performance is occasionally worse for experimental subjects than it is for controls (who have not seen the primes).

Priming questions provide a focus that tells the subject what aspects of the text are important. However, providing a focus can not be the entire explanation of the improvement due to priming questions. The location of priming questions is also important. Frase (1968) found that asking a question after each paragraph is superior to asking all the priming questions at the end of the passage. Frase (1967) also found that asking a priming question after two paragraphs was better than after one or four. If the effect of priming questions were merely to provide a focus, then the position of questions should not matter.

It seems that priming questions do more than provide focus; they force the subject to process the text in a certain way. This is clear from results of experiments where critical aspects of the text were directly highlighted (in a box) at study. Bruning's (1968) study found that this method of highlighting was not nearly as effective as forcing the subject to answer a question that caused that material to be reviewed in order to answer it. The conclusion that providing information may not be as effective as forcing the subject to retrieve it himself has been encountered before.

Bobrow and Bower (1969) found that providing the mnemonic to relate a pair of words was not nearly as effective as asking subjects to provide one themselves.

Frase (1967) suggested that priming questions cause the subject to review the relevant aspects of the passage. This review process probably involves more than merely stating the critical information in the question or answer since highlighting the critical information was so much less effective. Indeed, McGaw and Grotelueschen (1972) found that the information does not need to be directly tested. In their study, when a question reminded a subject of information present in the text, without either stating the information in the question or demanding it as an answer, subjects later recalled that fact better than a comparable fact unrelated to the priming question. The reason that priming questions in studies mentioned earlier for the most part only helped same questions and not new test questions is probably due to the nature of the overlap between questions. In fact, Reder (1976) and Watts and Anderson (1971) both conducted experiments in which priming questions helped some new test questions and not others, depending on the type of questions or the type of relation between priming and test questions.

It seems clear that merely asking a question will not produce improved performance. The question must force the subject to process relevant aspects of the text in "useful" ways. Watts and Anderson (1971), for example, found that subjects do better on a post-test about passages they read if they had been asked immediately after reading each passage a question that forced them to "integrate" the material (i.e., apply a principle in the passage to

a new example). On the other hand, subjects do even worse on the post-test if asked a low-level question (name the scientist associated with the principle) than if asked no question at all.

One explanation for this result involves the notion of generating useful elaborations. This notion is that the more extra processing one does that results in additional, related or redundant propositions, the better will be memory for the material processed. The importance of relevant elaborations or embellishments to memory has been argued elsewhere (Reder, 1976) and will be discussed in more depth in section 3. Elaborations provide redundancy in the memory structure. Redundancy can be viewed as a safe-guard against forgetting and an aid to fast retrieval. In the Watts and Anderson study, when subjects were asked no questions they probably had more time to generate useful elaborations than when they were forced to generate the name of the scientist which has nothing to do with the point of the principle. In other words, it is not enough to force people to think about the relevant information, they must generate semantically useful elaborations; if they think about aspects of the text that seem trivial and unimportant (e.g., when reading about how electricity works, thinking about the cost of electricity today), then they will not be able to answer questions about aspects of the passage that are deemed important.

Rothkopf (1972) conducted a study whose results are consistent with this view. In this study subjects read material at their own rate and they would read slower those passages for which they were given priming questions. They probably were attending (thinking, elaborating) more to the relevant information in primed passages. Anderson and Biddle (1975) also produced

results consistent with the notion of greater amount of attention producing better performance. In their experiments subjects were asked a priming question after reading the passage that could either test verbatim memory or gist memory. They found that a priming test of verbatim memory produced better subsequent performance on either a verbatim or paraphrase test than did a previous paraphrase memory test. If one assumes that one makes multiple redundant elaborations related to a statement during study, finding anyone will suffice for a gist recall test. However, to decide if something was said in that exact form requires a much more thorough search of one's memory representation. Perhaps verbatim search, being more time consuming, allows more elaborations to be rehearsed (strengthened) and more new ones to be generated.

The extra elaboration notion mentioned above seems consistent with most of the results thus far. There is some data, however, collected by Frase (1968, 1975) and Rothkopf and Kaplan (1972) that seems both unintuitive and difficult for the extra elaboration interpretation. Subjects were found to perform better on new test questions if the priming questions are more specific (e.g., "When was Jack born?") than when they are general (e.g., "When were the men in the paragraph born?"). Presumably, more general questions would cause more elaborate "general discussion" questions. On the other hand, at least some of the results used to argue for the superiority of specific questions are suspect. For example, one task used by Frase (1975) is unlike those tasks to which he wishes to generalize. Subjects' instructions were to either underline the names of all living creatures or of all land animals: the names to be underlined were either <u>category</u> <u>names</u>

or the animal's proper name in the story, e.g., "Jelly." These words were embedded in a story about a king who was in charge of naming all the animals. Frase was interested in the time to perform the task and the amount of recall of animal category names; he did not ask for recall of the story. This experiment seemed to be a verbal learning task and, in fact, was a replication of an earlier study (Frase & Kammann, 1974) that did not have subjects read a story. When subjects underlined every noun or every capitalized noun ("proper noun"), the task involved only low-level processing (Hyde & Jenkins, 1973; Craik & Lockhart, 1972) and was performed faster. In the conditions where subjects recalled more, they took more time and had to process words at a deeper level (consistent with the "extra thought processing" notion). The task is unlike a normal reading task.

The research discussed thus far has dealt principally with how to improve retention of prose material. Forcing readers to attend to those aspects of the passage that are deemed important results in better retention of that information. This research has focused less on how people come to comprehend a passage or why various techniques improve performance. The apparent inconsistency in results among some of this research probably stems from our inadequate understanding of the structure of passages, how content is represented and how it is processed and integrated with prior knowledge. Subsequent sections will review recent attempts to improve our understanding.

## 1.2 Research Using Additional Context as External Aids

The research of Bransford and Dooling illustrate the importance of a relevant conceptual structure within which to embed new information for

long-term retention of the input. Like the research discussed in the last section, this work explores the usefulness of adjunct aids. In this work, the adjunct aids have consisted of pictures, titles, topic phrases, etc. rather than questions. The goal of the research has been less concerned with techniques to improve recall (although recall has certainly been affected by the use of the aids), and has been more concerned with what affects comprehension broadly speaking. In many experiments the passages used were basically incomprehensible without such aids, while much clearer when aids were given in advance. The effectiveness of the aids has been explained in terms of Bartlett's notions of schemata. Bransford and Johnson (1973) and Bransford and McCarrell (1975) found that comprehension and ability to remember a passage were greatly improved when a descriptive title or picture was given prior to the passage. Without the title or picture, the passages were too abstract; the referents needed to understand the passage could only be inferred from the topic in one case and from the situation depicted in the picture in the other. In other words, comprehension depended on a schema or conceptual framework supplied by the experimenter to make the referents of the passage clear. Dooling and Lachman (1971) also presented passages that were much easier to recall if a topic (e.g., "The first space trip to the moon") was given before reading the passage. Giving the topic after having read the passage produced no better recall than not giving it at all (Bransford & Johnson, 1973; Dooling & Mullet, 1973). This indicates, I think, that if one cannot fit the input into a conceptual structure <u>immediately</u>, the phrases will not be transferred into long-term memory. Thus, the topic is of no use in

the after condition as the story is already forgotten. It is also the case that providing the subject with the wrong topic before reading an abstract passage makes recall worse than having no topic (Bransford & Johnson, 1973). Perhaps this is because the reader wastes processing time trying to integrate the input into the wrong structure without success (i.e., it does not "fit") or creates inappropriate elaborations. Had he been given no structure, the reader stood a better chance of deriving an appropriate one. These findings are consistent with Ausubel's suggestion that the reader profits from being given prior structure on which to map the incoming material.

Bransford and Franks (1971) argue that in "normal passages," when the referents are clear, a schematic representation of the input is constructed and only that schema is stored in memory. They showed that subjects were poor at discriminating exact sentences they heard from other sentences that were also derived from the "schema." Bransford and Johnson also showed that if the passages are concrete, (i.e., the referents are clear), prior knowledge of the topic does not aid comprehension or recall. This finding is not inconsistent with Ausubel's assertion that prior structure helps comprehension. He had found that an advance organizer's usefulness interacts with the reader's knowledge and the subject matter. In some "concrete" passages for which the reader knows little about the topic, prior exposure to the structure would undoubtedly improve comprehension and retention. Perhaps the term "concrete" should be defined with respect to each individual's semantic memory since various referents may not seem clear if the reader is unfamiliar with the topic.

Sulin and Dooling (1974) demonstrated the influence of prior knowledge in a different manner. They argued for the importance of thematic effects in retention of prose. When a subject read a passage that was said to be about a famous person, e.g., Hitler, they were less confident about rejecting in a recognition task foils that were dominant traits of the famous person than less thematically related foils. When subjects were told that the passage read was about a fictitious character, they rejected all foils with about the same confidence. Sulin and Dooling were able to find in one of their two experiments that the "thematicity" effect increased with delay, as predicted by Bartlett. However, most of these results are open to potential artifacts: The effect could be merely a response bias rather than a difference in encoding. To demonstrate that it is not a response bias, they should have contrasted their results with those from subjects who were not told the identity of the character (Hitler vs. Gerald Martin) until the time of test. Unlike a response bias model, a "schema" model would not predict a replication of their results in the control condition. This is because according to the schema theory, it is during input that one reorganizes, interpretes, etc., not at test. It should be pointed out that subjects rarely false-alarmed (i.e., said the foils had been presented); they just gave lower confidence ratings about related foils being new than about less related foils.

More recently, Dooling and Christiaansen (1977) performed the needed control experiment. The subject was told the identity either before reading the passage, immediately after reading the passage or immediately before test at one week. They found that subjects made many false alarms at a

week delay regardless of when they learned the "identity" of the famous cnaracter. The percentage of thematic false alarms was greatest when subjects were told immediately after reading the story. Neither the simple response-bias model nor the simple schema model predict the obtained pattern of data, that the most thematic false alarms occur when subjects are informed right after reading the story. However, a reconstructive model plus a type of response bias model together could handle the data. The model would posit that if the subject learns the identity of the fictitious character immediately after reading, then the subject will go back over the story, mentally, reworking it to make it consistent with the "true" identity. This causes more thematically related inferences to be stored erroneously than when the subject knows while reading the story that the passage is about a famous character. A response bias must also be operating (a pre-disposition to respond in a manner consistent with knowledge about the famous character) to account for the greater number of false alarms to thematic foils when told the character's identity immediately before test.

One interpretation of Sulin and Dooling's result is that subjects need not store the related facts in long-term memory since the result is partially a response bias. This view is supported by data of Frase (1970, 1971). When subjects were asked to rate the validity of statements with respect to a passage, no difference in ratings of inferences and assertions were found. Frase did find differences in recall, however. Very few inferences tended to be produced. This might be because the inferences were not made until subjects were asked to judge their validity; or they might

have been made, but tagged as inferences and not considered appropriate for recall. The former seems more likely because when subjects were forced to draw the inference to answer another question, the incidence of the inference's recall increased.

In conclusion, the findings of Bransford and Dooling discussed earlier, like Bartlett's, provide evidence that context and general knowledge affect memory for prose. They go beyond Bartlett's work by showing when "distortion" of memory can occur (i.e., immediately after reading as well as during reading) and when the requisite knowledge is needed for comprehension and retention (i.e., it must precede reading). Like Bartlett, however, the investigators offer few insights as to what the structure of schemas might be like or the mechanisms that allow schemas to affect comprehension. Anderson and Biddle (1975) remark with respect to the questioning literature reviewed in the last section that "We do not need another demonstration that adjunct questions 'work' . . . [we need to know] . . . why they work . . . (p. 108)." The same basic comment could be made with respect to these demonstrations. An explanation for these findings will be given in section 4.

### 1.3 Investigations that Test the Generalizability of Verbal Learning Results to Connected Discourse

The comprehension and retention of prose has also been investigated from the perspective of how much generalization there will be from research using smaller units of analysis to this level of analysis. It seems reasonable to determine whether earlier work is applicable to more natural settings and to try to make use of the products of past research.

Proactive inhibition and retroactive inhibition are two classic findings from learning research (see Postman, 1961). Proactive inhibition occurs when subjects first learn a list of A-B pairs, (i.e., must recall the 'B' responses to the 'A' stimuli) and are then taught a list of A-C pairs in which new responses are paired with the old stimuli. Although subjects can generate the C responses to the same A stimuli, they forget the C responses faster than a control group that learned an A-B list and then a D-C list. Retroactive inhibition is the phenomenon that a delayed recall of the B responses to the A stimuli is worse in the A-B, A-C condition than in the control A-B, D-C condition. Positive transfer is another related phenomenon in which learning A-B' is facilitated (learned quicker) than A-C or D-C when subjects learned A-B first (the B and B' responses are very similar.)

Interference and positive transfer have both been obtained in recall of one passage when a second passage has also been studied between its presentation and test. Whether interference or positive transfer is found depends on the nature of the overlap of the two (Anderson & Carter, 1972; Bower, 1974; Myrow & Anderson, 1972; Walker, 1974). Anderson and Carter found retroactive inhibition in the subjects' ability to recall the subject of a sentence to the predicate cue. This occurred when a different sentence containing a paraphrase of the predicate and an unrelated subject were presented for study between the original sentence's study and test. Walker obtained retroactive inhibition for recall to a first paragraph after reading a second related one.

Bower (1974) investigated which aspects of passage overlap lead to facilitation and which lead to interference. He found that recall of the macrostructure (i.e., the general organization and topic of the information covered) improved if a different passage with the same conceptual macrostructure was presented between the critical paragraph's presentation and test. Recall of this macrostructure was improved independent of whether the intervening passage with the same conceptual macrostructure had the same microstructure (details) or a different microstructure; recall of the microstructure, however, was inferior if the microstructures differed (and the macrostructures were the same).

Organizational factors have been shown to be of importance to recall of prose just as they have been in free recall of word lists. Not surprisingly, Montague and Carter (1973) found subjects recall more of a passage which maintains a correct syntactic order than one in which words are presented in random order. Even in non-scrambled texts, the input organization of concepts and their associated attributes can affect the amount recalled (Frase, 1969, 1975; Myers, Pezdek & Coulson, 1973); subjects tend to organize recall in a manner consistent with the organization of input, i.e., by name or by attribute, (Perlmutter & Royer, 1973). Serial position effects have been found with prose passages (Deese & Kaufman, 1957; Frase, 1969; Kirscher, 1971) while others (Richardson & Voss, 1960) have failed to find the effect. Meyer and McConkie (1973) claim to have resolved the apparent inconsistency. They separated serial position from structural importance, using their representational system (to be described later), and found recall is better predicted by the latter. I suspect, however,

that if one can control structural importance, one would still find an effect of serial position since it is such a robust finding with other materials.

The research on prose which extends the work typically associated with word pairs and sentences argues against the criticism that investigations with smaller units of analysis will not generalize to normal text. However, the conclusions one should draw from the results are not clear-cut. Interference, positive transfer and organizational effects are only achieved with paragraphs when much care is taken in construction of the paragraphs so that they resemble each other in specific respects. Myrow and Anderson (1972) acknowledge that their obtained results demonstrating retroactive inhibition between passages are unlikely to occur in the classroom because typically one does not encounter sets of passages that match (and mis-match) in such specific ways.

It would be worthwhile to continue investigating the degree to which effects found in verbal learning generalize to prose processing, but only within the framework of a theory of memory and comprehension. A basic complaint can be lodged against much of the research discussed in that it concerns the discovery of the conditions under which interference effects generalize to prose processing, but not the explanation of why and how they generalize. The research would seem more important if memory mechanisms that produce RI and PI in both verbal learning and prose paradigms were proposed. Whether or not one obtains serial position effects with prose is less interesting than knowing why one should or should not obtain the same effect as that obtained with word lists. Anderson and Bower (1973) propose the only theory that deals with these issues.

## 1.4 Discussion

A number of results can be distilled from the preceding studies. The extent to which a passage is comprehended or retained is not solely a function of the text or the reader. Rather, a number of experimental manipulations can interact with text and reader to affect performance. The type of questions asked about the passage, whether background information is presented, how easily subjects can map referents onto familiar concepts all affect performance. In some circumstances it doesn't matter whether the questions are asked before or after reading the passage; in other cases better performance (ability to answer questions) occurs when questions are asked prior to reading; in still other situations, performance is best if questions are asked after the passage has been read. When passages are sufficiently vague, they can only be understood if a title or picture is given that indicates the setting.

Generalizations concerning effects of context, priming questions, interference or transfer (e.g., when will manipulations have the greatest effect), are difficult to specify. Which manipulations are effective at which times could be predicted by a model for prose comprehension. Apparent inconsistencies in results could perhaps be understood within a theoretical framework.

One of the requirements for such a comprehension model is a good representation of text. The research discussed thus far was not concerned with developing one.

## 2. REPRESENTATIONS OF TEXT STRUCTURE

Many theorists concerned with finding an adequate representation of a prose passage maintain that doing so is a necessary prerequisite to other research on prose comprehension. Meyer (1975) has argued that structural variables can influence the learning and retention of items in a word list and that structure probably also affects retention of prose passages. She points out that one can ameliorate problems such as measuring the similarity and differences among passages and scoring the recall protocol from a passage by the use of a structural analysis of prose. Others (e.g., Crothers, 1972; Frederiksen, 1975a) have argued that theorists will be able to induce the processes involved in text comprehension by comparing memory protocols of a passage with a representation of it. Whether or not these theories of representation are useful for all aspects of research on text understanding remains to be seen. For example, it is not obvious that investigations concerned with the optimal position for a priming question to facilitate the acquisition of specific facts hinge on a structural analysis of the passage to be read. Whether structural variables affect priming facilitation is an empirical question.

There have been a number of approaches to the problem of specifying the representation of a text or a story. The first sub-section to follow concerns specifying logical relations within a story or text; other representations are concerned with all relations found in a text or a story. The third sub-section describes theories that have rules for generating a story. Each representational theory is concerned with either expository text or stories, rarely both.

## 2.1 Representations of Logical Relations

### 2.1.1 Dawes

The earliest work to be concerned with semantic representations of prose was that of Dawes (1964, 1966). Dawes was interested in investigating what cognitive distortions people tend to make, especially distortions of relations asserted from the material read. He looked at set relations in stories and tested the idea that a person's memory for set relations would "simplify" over time, specifically that disjunctive relations would become nested relations (see Figure 1). Inclusion, exclusion and identity are all defined as nested, while overlapping sets are disjunctive. Going from disjunctive to nested is overgeneralizing (from "some, but not all" to either "all" or "no") and the process reduces the number of categories in the environment by deleting that part of set x not nested in set y. He did find distortions and a tendency to simplify, but contrary to Bartlett's thesis, he failed to find increasing simplicity of recall at greater delays. His results can be taken as evidence that the phenomenon occurs at input or output and has nothing to do with forgetting.

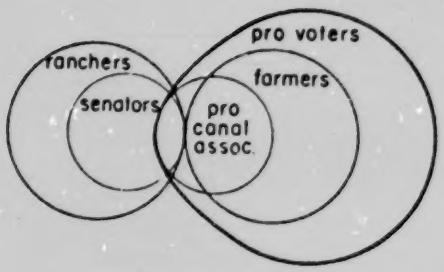



Figure 1: Venn diagram of Dawes' "Circle Island" story.

NESTED relations

inclusion

all senators are ranchers

exclusion

no senators are farmers

identity

none in figure, but all ranchers are wealthy and all wealthy people are ranchers

DISJUNCTIVE relations

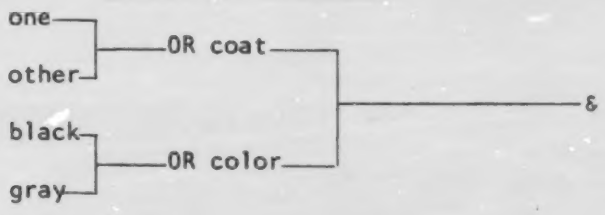
overlapping sets

Some senators belong to pro-canal association

### 2.1.2 Crothers

Crothers' (1972) work is similar to Dawes' in that it too concentrates on the logical relations in a passage. Crothers measured the extent to which prose recall would be predicted by his representation of the passage. His representation only uses logical relations and a semantic hierarchy which classifies concepts under their superordinates (that need not have been mentioned in the test), rather than a richer relational system. The original organization provided by the author is lost. The semantic

hierarchies are connected together by the logical relations IS, WHY (because of), OR, AND, and IFF. An example subgraph for the phrase, "One coat was black and the other was gray" is given below:




Crothers failed to obtain the prediction that concepts at a higher level in the hierarchy (defined by the semantic memory network of Collins & Quillian, 1969) were more likely to be recalled. Overall, his conceptual representation did a poor job of accounting for recall of passages. His idea of trying to represent the story as it would be in the reader's memory was a good idea, but his view of the reader's memory seems too impoverished: He implied that one's semantic memory only notes set inclusion inferences, e.g., a dog is an animal and a cat is an animal. However, it seems other types of relations and idiosyncratic information is brought to bear in understanding a story. Crothers' representation seems overly influenced by Dawes' analyses of set relations in stories and Collins and Quillian's theory of the organization of memory. Crothers may have found it easy to represent his passages using his representation, thereby lending initial credibility to his notions. His passages however, were extremely descriptive (e.g., describing two kinds of nebulae) and thus could easily be represented by set relations. Most passages are not so descriptive in character and would have made more apparent the deficiencies of his theory for representing text.

### 2.1.3 Frederiksen

Frederiksen (1972) had a model similar to Crother's in that Frederiksen's semantic structure graph used both set relations and subject transformations (i.e., implication relations). An example structure is given in Figure 2. Frederiksen performed an experiment using Dawes' Circle Island story and he has analyzed it in a number of ways (Frederiksen, 1972, 1975a, 1975b). Subjects read the story four times: two of the three groups recalled after each reading, and the other group was given a surprise recall after the last reading of the story. They all recalled once more a week later. The uninformed group and one of the two informed groups also had an orienting task of a problem solving nature, related to the passage, to work on after each reading. The dependent measure was recall, and Frederiksen developed a scoring system that noted whether concepts were overspecified (pseudo-discriminated), e.g., colonel for officer, or underspecified (overgeneralized) e.g., military man for officer. He looked at how orienting tasks affected the specification of concepts and set relations and how this recall changed as a function of successive recalls.

The group given a surprise test recalled less than the informed groups. This group and the other problem solving group produced more elaborations (statements not derivable "inferentially" from the passage), overgeneralizations and inferred relations than the group just asked to recall. Both informed recall groups gave more "veridical" output than the uninformed group, and the two informed groups did not differ in the amount of veridical recall. The non-problem solving group gave more pseudo-discriminations.

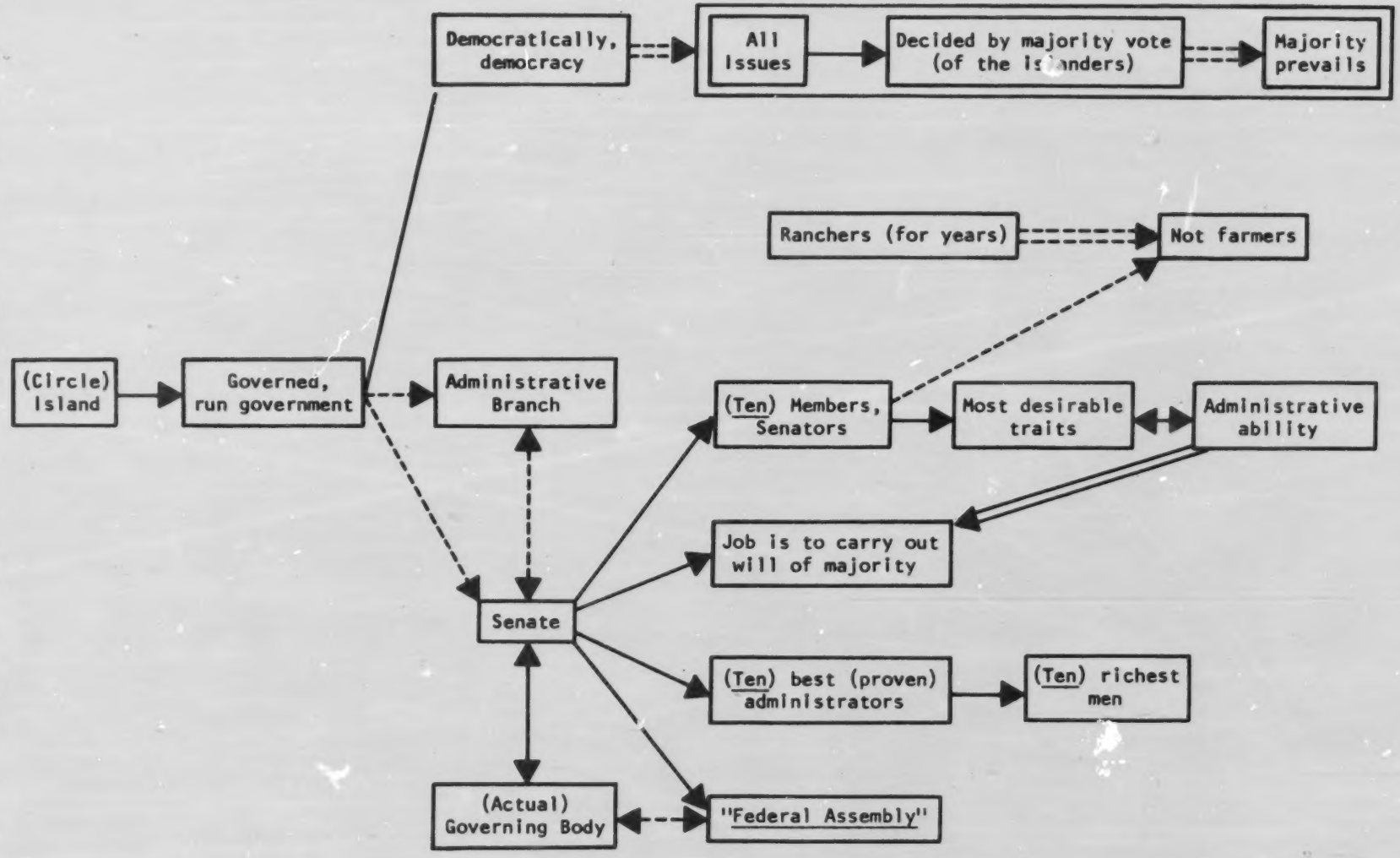



Figure 2. Frederiksen's Semantic Structure Graph of first paragraph of the Circle Island Story.

Frederiksen believed that the problem solvers had a greater processing load and would therefore have to encode the input more simply. He argued that the inferences produced in recall are products of simpler coding of the input. With these two assumptions he explained the greater amount of inferential recall of the problem solvers. However, this argument seems wrong for several reasons: First, drawing inferences from the input should require more processing, not less. Second, the amount of veridical recall did not negatively correlate with inferential or derived recall. Therefore, inferences could not result from a "greater processing load" in the problem solving condition. Perhaps instead, "semantically relevant" memory errors occur more frequently for subjects given a meaningful orienting task to direct their prose processing because meaningful processing involves building a representation of the gist and ignoring the exact input phrases. A greater processing load theory would imply, on the other hand, that these subjects should lose more of the information irrelevant to the task. There was no evidence of this. One explanation of why problem solvers produced more elaborations not derivable from the story is that they import more world knowledge in order to solve the problems. These imported facts needed for solving problems may become confused with the input.

In analyzing the errors over successive recalls, Frederiksen (1975a, 1975b) found more inferred concepts and more overgeneralizations on each output trial, but successive recalls lead to fewer elaborations. He considers this evidence that generalizations occur during input (reading the test) and elaborations are generated during recall. Presumably elaborations

are edited during the input processing and generalizations are not, because generalizations are not corrected on successive input trials. There is another interpretation for this result, however. The more often one recalls, the more boring the procedure becomes and hence, the more shorthand expressions one looks for. The group that only recalled once may have put down more elaborations because they had little veridical recall and wanted to say as much as possible. The groups who were then recalling for the fourth time may have omitted the elaborations to get through the task as quickly as possible. The inferences and overgeneralizations may have the property of making the output shorter; any shortcut that expressed the same idea would be desirable.

To test this alternative explanation, however, a control group would be needed that Frederiksen did not provide: a group that read the story once, but recalled four times. His theory would predict no additional inferences after the first recall, but more elaborations on each successive output. The other view predicts the opposite outcome.

One problem with Frederiksen's theoretical claims is that they are all based on one study using one story, Circle Island. Dawes wrote Circle Island to investigate errors of logical relations. Frederiksen's representation for stories, like Crothers' will probably not fit most stories. To my knowledge, this representation has not been tried out on a variety of passages.

## 2.2 Representations Concerned with More Complex Relations

### 2.2.1 Frederiksen

Frederiksen's (1975c) model has become more detailed and takes into account far more aspects of language than his earlier representations.

His representation involves two graphs, the semantic structure graph and the logical structure graph. The first is reminiscent of Fillmore's (1968) case grammar for sentences. He diagrams stative relations, case relations, attributive relations, etc. into semantic hierarchies or networks, making fine, semantic feature distinctions. The logical structure graph is not hierarchical; propositions from the semantic structure graph are nodes connected together by the logical, causal and algebraic relations. The logical relations and causal relations include conjunction, disjunction, logical implication, and material biconditional. The algebraic relations are order (transitive) relations and proximity relations. He can also represent relative time and location.

Frederiksen's representation is one of the few to include inferences. That makes it a more realistic representation of what the long-term memory structure might be like as a result of comprehending a passage. Unfortunately, his representations seem complex at the wrong "level" of understanding of a passage. The strategy of decomposing concepts and relations into semantic primitives does not seem viable as a representation for prose. One could argue (see discussion of Kintsch, below) that people do not decompose words into their semantic primitives during comprehension. Even if they do, a representation of an entire text should focus on the many higher order, complex relations that express the message of the passage, not the lower-level complexities. Frederiksen's system seems devised for representing the relations within a sentence, not among sentences. The network has capacities for connecting propositions referentially and temporally, but apart from these, there is little in his system that seems concerned with the

text level as opposed to the sentence level. Textual aspects which he does not represent include the overall organization of the text of the plot and inferences and elaborations that connect lines of input.

### 2.2.2 Kintsch

Kintsch (1974) has also developed a formal representation for texts and has performed numerous experiments to test the validity of this theory. The fundamental unit in his analysis is the proposition. A text is represented as an ordered list of propositions. He avoids network representations such as those described earlier, due to their impracticality for passages of any length or complexity. Networks are messy for single propositions, and they rapidly become unwieldy when expressing multiple propositions. A sentence may consist of one or more simple propositions, each proposition consisting in turn of a relation and a set of arguments. The representation has the flavor of predicate calculus. Verbs and adjectives are relations (attributes), and nouns and other propositions are its arguments. Propositions are connected together by means of a repetition rule, in which one proposition is referred to by a "superordinate" proposition that contains this subordinate as an argument. This forms a hierarchical structure. Propositions are also connected by sharing the same arguments or concepts, e.g., an individual of a story can be in multiple propositions. The first proposition that uses the shared argument is considered the superordinate of the proposition that contains the repeated argument. Propositions can also be connected by an ampersand (&) when they share a common argument. For example, Kintsch (1974, p. 63) represents "John broke his leg yesterday"

as: ((BREAK, JOHN, LEG) = α) & (TIME, α, YESTERDAY). Kintsch claims his hierarchical organization of text has predictive value for recall.

Kintsch does not include any inferences in his examples of text representations, yet his theory does include rules for generating some inferences. It is not clear from his rules what inferences could or could not be generated, nor which inferences are likely to be stored as part of comprehension.

Kintsch reports numerous experiments in support of his theory (e.g., Kintsch & Keenan, 1973; Kintsch, McKoon & Keenan, 1974; Kintsch & Monk, 1972). He admits that these experiments support a class of representations, of which his is one; some studies, though, (e.g., Kintsch, 1974, ch. 7) were designed to distinguish between his propositional representation and other extant proposition representations. Below a few of his studies will be described in varying detail.

In order to determine whether subjects draw and store inferences while reading, Kintsch, McKoon, and Keenan (Kintsch, 1974, ch. 8) looked at reading times for paragraphs that either stated or implied facts that followed from the text and looked at time to judge if these facts were true of the text. The "level" of difficulty of the inference was manipulated in one study; the delay from study at which the question was asked was varied in a series of experiments. Table 1 gives the example used in Kintsch's book for the different levels of inference. Kintsch et al. in one condition actually presented the inference as part of the text (explicit condition); in another condition the statement was not presented, but could be readily inferred (implicit condition). They found no difference in reading times that depended on

Table 1

Level 1
A strong hand was needed to restrain the dog. The animal's instincts had been aroused by the sight of the fleeting deer.
(The dog was an animal.)

Level 2
A burning cigarette was carelessly discarded. The fire destroyed many acres of virgin forest.
(A discarded cigarette started a fire.)

Level 3
Police are hunting a man in hiding. The wife of Bob Birch disclosed illegal business practices in an interview on Saturday.
(Bob Birch is the man who is hiding.)

whether the statement was explicit or implicit, and no systematic effect on reading time due to level of inference. Time to verify a fact did not depend on level of inference either. Subjects were faster to verify statements that were explicit in the text immediately and a half minute later, but the difference in reaction time to verify explicit versus implicit statements was not significant 20 minutes or 48 hours later. They concluded that there is an immediate surface structure advantage for the presented condition, but over time this superficial difference in memory representations evaporates, i.e., subjects have stored the inference in the implicit condition and the underlying representations in memory are the same. One might question this interpretation, however, since the appropriate statistical test was not conducted.[5]

Kintsch and Monk (1972; Kintsch, 1974, ch. 5) argued that ideas in a text are represented in propositional format which does not depend on the surface structure complexity of the input. Two versions of a passage were

constructed, one simple in structure and one complex, although both were presumed to convey the same ideas. Subjects read one of the two versions. Reading time was of interest as was time to judge that an inference from the story was valid. They found that reading times were longer for complex passages, but that latencies to judge inferences did not vary as a function of complexity. These findings are consistent with their notion that it takes longer to comprehend complex statements and encode them into a propositional format; once stored in their canonical representation, however, the time to make the inference should not reflect surface structure. Kintsch also suggests that recall can be predicted by level in the hierarchy, superordinate propositions being recalled more frequently. Kintsch and Keenan (1973) found this trend. McKoon (1977) found subjects are faster and more accurate to verify topic information than details and that the discrepancy between topic and details grows with delay.

A number of theorists (e.g., Rumelhart, Lindsay & Norman, 1972; Schank, 1973) and Kintsch at one time (Kintsch, 1972) have claimed that words are decomposed into their semantic primitives during comprehension. Kintsch (1974, ch. 11) reports a series of experiments testing the validity of this notion. Using the derivations of generative semanticists (e.g., Lakoff, 1970), Kintsch determined which words are more complex, i.e., have the longest derivation to unpack. These words, if decomposed when parsed, should be more difficult to process. Examples of complex words are kindness, prosecutor, accuse and persuade; simple words include environment, orchestra, relate and save. The word "persuade" should decompose to "cause y to intend to do x." In one experiment, Kintsch had subjects complete sentences (make

a sentence out of a phrase) while varying the complexity of a word in the presented phrase. Length, imagery, and frequency were controlled. Latency to initiate writing in a sentence completion task was not a function of complexity. Phoneme monitoring tasks did not reflect this factor either. Failing to obtain a difference in sentence processing, Kintsch looked at recall of sentences with words of differing complexity, but found no effect here either. He concluded from this series of studies that one does not typically decompose words during comprehension. He notes, too, that it is unlikely that we would have developed single words to express complex ideas if the words had to be "unpacked" each time they were heard.

In some respects, Kintsch's (1974) theory is rather similar to Frederiksen's latest (1975c) theory. Fredriksen gives more emphasis to a low-level (sentence) representation while Kintsch's hierarchical structure seems somewhat better suited to text organization. Certainly a linear notation is more tractable for text representation than is a network structure, although formally, there is no difference. Both theorists include mechanisms for cross-proposition inferences. However, they both seem more concerned with representing sentences within a passage than the ideas of the passage itself, i.e., the inferences in both systems seem to be narrow in scope. They are not as low-level as those of Dawes, Crothers and the earlier work of Frederiksen, but the inferences tend to focus on such problems as causation and reference. Table 1 illustrates the levels of inference used in Kintsch's experiment. They give an idea of the range of inferences he considers.

The inferences one would like to see in their representations or models for comprehension are those derived from knowledge about the passage as a whole, e.g., inferences about characters' motivation, goals, and responses to events. Elaborations about spatial locations, instantiation of physical details of people and objects are probably made during comprehension, too. Perhaps Kintsch and Frederiksen consider the elaborations suggested above to be idiosyncratic, and therefore, impossible to represent. Kintsch does state that some inferences are made during comprehension, and in a chapter (Kintsch, 1974, ch. 4) in which he offers a theoretical explanation for Encoding Specificity (Tulving & Thomson, 1973), he postulates the use of elaborations to modify the memory structure of "episodic events" (Tulving, 1972). Unfortunately, his notion of elaborations here seems restricted to phonemic information, superset and property information.

2.2.3 Meyer

Meyer (1975), like Kintsch, has a hierarchical rather than a network representation. Her theory adapted from Grimes (1975) has a greater resemblance to Fillmore's (1968) grammar than do other theories. The similarity between her model and Fillmore's is in the emphasis of the role or case terms to represent the relation between the verb (predicate) and the other words (arguments) in the sentence. (Use of Fillmore's case grammar has also been adopted to a lesser extent by Rumelhart, Lindsay & Norman, 1972; Kintsch, 1972, 1974; and Frederiksen, 1972, 1975c). Meyer lists nine cases or roles quite similar to Fillmore's, which she calls "lexical predicates". Unlike Fillmore (but very much like Grimes), she has "rhetorical predicates" at a level above the lexical predicates. These higher level predicates give

prose its organization. Neither Frederiksen's nor Kintsch's representations (discussed above) have such a well-developed, higher-level organization.

The representation Meyer uses resembles an outline of a passage. Propositions are indented according to their importance. Unlike an outline, however, every idea is included and relationships are indicated between branches at the same level and at differing levels. These relations are the rhetorical predicates, and they fall into three categories: (a) paratactic relations, which have at least two arguments of equal weight (e.g., RESPONSE: problems and solution); (b) hypotactic relations in which the arguments are unequal (e.g., SPECIFIC: problem and the details of the problem); and (c) neutral relations in which a collection of arguments can have equal weight or not (e.g., ATTRIBUTION).

Meyer has attempted to validate her representation empirically with a recall experiment. The experiment was designed to show that the probability of recall of a sentence was a function of its "height" in the content structure (i.e., its importance indicated by degree of indentation). Meyer constructed two passages with the identical target paragraph embedded in each. The paragraph was placed high in content structure in one passage and low in the other, but its serial position was the same. She found that the two passages as a whole were recalled equally well but that the recall of the target paragraph differed as a function of "height." The differences in recall between the two paragraphs increased with a week's delay.

After the free recall task at a week's delay, subjects were given a cued recall test. Cued recall in her experiment meant giving subjects all the

content words from the passage in any unsystematic (scrambled) order and asking them to arrange the words in the correct order. Even in this task, the advantage of the paragraph high in the content structure was maintained over the paragraph low in the content structure. From these results, Meyer concluded that the difference in memory representation for "low" and "high" paragraphs can not just be a function of differential accessibility. She argues that if the paragraphs were stored equally well, any differences due to accessibility would be avoided by the cued recall methodology. Meyer's task, however, seems sufficiently different from the standard cued recall task that differential accessibility still seems a possible explanation. People have great difficulty solving anagrams, yet few would doubt that these puzzle solvers have the word (whose letters are presented in scrambled order) stored in memory. It may well be in her task that subjects tried to retrieve aspects of the story and then looked for the words to put in the appropriate slots of the retrieved aspects.

Meyer found that the frequency of recall of an idea unit high in the paragraph structure was determined by the unit's functional relationships to other units, rather than its content per se. This was not true, however, for idea units that were low in the content structure. In other words, two different passages with the same structure of specific relations, but with different content, would yield a high positive correlation of probability of an item's recall from the same point in the structure, so long as that point is a "high" point; items from the same point low in the structure are not correlated with respect to recall. In these cases, the words recalled are quite idiosyncratic to story and subject.

From this result and ones mentioned earlier, Meyer concluded that elements "high" in the passage are given more processing time than elements lower in the passage. Perhaps all higher-level material is attended to, and with the remaining "processing-time" the subject attends to that low-level information which "catches his eye" (i.e., those units that seem inherently interesting). This would explain why, at low levels, structure does not predict recall. The mechanism Meyer suggests to explain differential retention is additional rehearsal. It has been shown, though, (Bjork & Jongeward, 1973) that additional rehearsal, per se, does not improve recall. One possible process might involve generating inferences and elaborations which draw upon previous knowledge of related information; the information higher in the structure would be elaborated more often than low-level information.

There is an explanation for her differences in recall which does not posit differential processing at input. Low-level information may have been stored equally well, but not recalled as well due to output interference (Roediger, 1973, 1974). The phenomenon of output interference is that, other things equal, the earlier something is tested, the better it is recalled; recalling some things has a damaging effect on recalling other things (see Crowder, 1976, for a discussion). In other words, the subject would recall aspects of the story high in the content structure first because they are more important or essential. Things that are low in priority are interfered with because of this bias in order of recall.

Meyer's representations do not include any inferences or elaborations, not even ones necessary to connect various propositions of different sentences.

Clearly it is not her goal to represent the structure of prose in the reader's head; if that were her goal, she would not include every detail from the passage and she would include many inferences and details people integrate into prose when they read it. Since this is not the intent of her representation, one might ask what her structure is to represent. The fact that the structure looks like an outline strongly suggests that she represents what the author wanted to say, not what the reader absorbs. Her representation can be used to predict recall and as a metric to compare passages. While these aspects are useful, representing the author's intended message is less interesting to psychologists than is representing the reader's comprehension of it. Meyer's representation might have been more interesting if she had indicated how the author's intended meaning is transformed into prose and then what portion of the original meaning is represented in the "comprehender's" long-term memory. This might have predicted recall still better.

There are a few other criticisms that can be made: For propositions represented "high" in her hierarchy, Meyer obtains systematic recall effects such that propositions with one structural relation are recalled better than propositions with other relations at the same level. Although at high levels, she found recall highly correlated with type of relation, she offers no explanation for why one relation should be better recalled than another. Meyer has only tried to apply her representation to passages that consist largely of factual statements, much like Crothers'. The representation does not seem suitable to a story which involves a plot with motivations, character reactions to events, etc.

Meyer's thesis has been a major contribution to the study of prose processing and retention. She has developed the first representation that actually seems concerned with passage level information rather than sentence level material. Her representation has been empirically validated and will be a useful tool in the design of future recall experiments using prose. For example, it is possible that various manipulations, such as priming questions, which affect a subject's ability to answer questions, will interact with the information's "height" in the passage representation. Presumably an important point will be less affected by a focusing manipulation than would a detail. Another use of Meyer's representation is to insure that two passages are equally complex. This would be important if one wishes to use a within-subject design in which subjects were to process the two passages differently.

## 2.3 Formal Models of Text Structure

There have been a number of attempts to identify the general organizational structures in prose (e.g., Colby, 1973; Lakoff, 1972; Mandler & Johnson, 1977; Propp, 1958; Rumelhart, 1975; Thorndyke, 1977; van Dijk, 1977). Developing a grammar for discourse has some of the same attractions that work on a grammar for sentences has, viz., the task is well defined, and it is therefore quite easy to determine when passages fail to fit the grammar. Although Meyer has been concerned with text structure, it is not clear that her work belongs in this group because her theory is less formal: She does not have a grammar to generate a passage nor formal procedures for selecting which propositions fit under which rhetorical predicates.

2.3.1 Theory of Discourse and Macro-structures

Van Dijk (1975, 1977; Kintsch & van Dijk, 1975; van Dijk & Kintsch, 1976) has been concerned with a general theory of discourse. Van Dijk has incorporated Kintsch's work discussed above (Kintsch, 1974) into his theory of semantic representations for sentences and sequences of sentences. This representation is called the micro-structure of the passage. Van Dijk has also developed a theory of semantic representation for global discourse structures. The global structures are called the macro-structures of the passage. Macro-rules allow one to relate the micro-structures to the macro-structures. Macro-structures are the main aspects of a text that are stored during comprehension.

In van Dijk and Kintsch's theory, discourse can be expressed in terms of super structures which are even higher level than macro-structures. They have limited their work to the narrative structure, one type of superstructure. A narrative structure is not itself linguistic, i.e., the structure can be expressed through pictures, etc. When it is expressed linguistically, it is called a story. A simplified narrative grammar is given in Table 2. Narrative rules do not seem as useful as macro-structures in that not all linguistic passages fit the grammar (e.g., text books, diaries, journal articles will not fit the rules). Because of this lack of generality, van Dijk has chosen not to devote much attention to this generative grammar for stories and rather concentrates on macro-structure.

One reason van Dijk devised macro-rules and investigated macro-structures may have been due to his experimental results: He asked subjects to read a story and later recall it. Other subjects were asked to read and summarize

the same passages. He found a great deal of consistency in the recalls and in the summaries. The common propositions found in all subjects' recall formed a sort of story itself and it was quite close to the summaries from the other subjects. These consistently recalled propositions are considered to comprise the macro-structure. Propositions most likely to be recalled either introduce main characters, give the characters' goals, or describe the actions leading to these goals and results from the attempt. Propositions concerned with settings, mental actions, etc. tend to be ignored.

Table 2

| | | |
|---|---|---|
| NARRATIVE | → | ACCOUNT + MORAL |
| ACCOUNT | → | SETTING + EPISODE |
| EPISODE | → | HAPPENING + EVALUATION |
| HAPPENING | → | COMPLICATION + RESOLUTION |

In van Dijk's analysis, the micro-structure usually entails the macro-structure. He derives three macro-rules: generalization, deletion and construction. These are posited to transform a text base into the core or essential macro-text base. The rule of generalization allows one to summarize a sequence of statements. There exist different levels of generality of abstraction entailed by the level of explicit text. The level of abstraction that is taken is the least general which also defines the smallest superset. The abstraction process stops when a superclass would not shorten the number of propositions. For example, statements in a text such as "Peter has a dog,

and three cats. He has a parakeet, etc." would be reduced to "Peter has pets," not "Peter has animals" or "Peter has things." A double application of the generalization rule can occur in a case where both the subjects and predicates of sentences are generalized: "Mommy was baking cakes, Daddy was gardening while their son washed the car and their daughter sewed a dress," can be summarized as "The family was working on household chores." Inferences of this type can be said to be true by definition.

Inferences using the rule of construction, on the other hand, are probably true but not necessarily true. Construction refers to finding a summary term for a set of actions. With this rule, the summary term is not entailed by the actions, but if all the required actions are stated, one infers the event. For example, "Peter laid foundations, built walls, built a roof..." becomes "Peter built a house." Whether one infers a particular summary depends upon the independent likelihood of that interpretation. If one reads that John went to the train station, a summary might be that he is going to take a trip, although this interpretation does not have to be true.

The other important macro-rule, the rule of deletion does not involve inferences. It follows the principle of irrelevancy, i.e., delete propositions that have no consequence, that are merely details. For example, one may delete from "Mary hit a blue ball that broke a window," the detail that the ball was blue. Other types of deletions include omitting extra specification of an act. The phrases "John went to London yesterday. He went to the airport, purchased a ticket and got on a plane bound for London..." are rewritten as "John went to London yesterday." One often deletes preconditions, normal conditions and some normal consequences, so that "John went to

the bank to cash a check" might be sufficient to assume that the bank was open and that he received money and not worry how he got there.

Van Dijk and Kintsch have continued to explore the usefulness of the notion of macro-structures (van Dijk & Kintsch, 1976; Kintsch & van Dijk, 1975). They are not interested in the role of inferences, per se, but wish to demonstrate the importance of macro-structures to comprehension. However, their results also demonstrate how the reader uses his conceptual structures and world knowledge to infer what must be the appropriate organization of the text. Subjects read passages with the intra-paragraph organization intact, but the inter-paragraph organization scrambled for half of the readers. They felt that this essentially scrambled macro-structures without disturbing micro-structures. Although subjects took much longer to read stories with distorted macro-structures, the time to generate summaries was not significantly longer for the scrambled group. The summaries made by subjects who read scrambled passages could not be distinguished (by other subjects) from summaries of subjects who read normal passages. These results are consistent with the notion that subjects work harder to assemble or infer the macro-structure when the paragraphs have been scrambled (hence, the longer reading time), but that once this is done, the representation will be the same as that for an unscrambled passage. This study resembles that of Kintsch and Monk described earlier.

In a related experiment with children, Kintsch has found a great deal more consistency in recall of the important (macro-structure) propositions of a story when the story was presented with the macro-structure intact.

Both this study and the one just described above, were taken as evidence of the existence of macro-structures. Although the notion of macro-structures is a useful one and the experiments are suggestive, one should ask how many of the possible set of outcomes from their experiments would be considered as consistent with the notion of "macro-structures." Note that in the case of the children's data, a failure to make use of the scrambled order was considered as evidence for the existence of macro-structures in children; for adults, the ability to overcome the scrambled order and produce a summary indistinguishable from those of normal passages is considered as evidence for macro-structures, too.

Greeno and Noreen (1974) demonstrated the importance of structure to comprehension in a study similar to those of van Dijk and Kintsch. They measured the time to read each sentence of a story and varied the ease with which the organization of the passage containing the sentences could be inferred. They gave subjects one of two presentation sequences. In either case, the information could be represented as a set of linked hierarchies. In the difficult condition, the information was presented to the subject in bottom-up order; in the other condition, the subject read the sentences top-down. It took subjects only two-thirds as long to read low-level sentences in the hierarchy if they were preceded by sentences above them in the hierarchy. There was not a comparable savings for high-level sentences preceded by low-level ones. Greeno and Noreen suggest that this result occurred because relationships involved in low-level sentences are consistent with the more general ones given in the higher-level statements. Subjects can process these lower-level sentences faster

because part of the conceptual structure has already been built. When the order is reversed, the procedure is less efficient. As would be expected by these notions, a sentence inconsistent with the structure took longer to process only when it was read after high-level statements. When the same sentence was read before the structure was obvious, the subsequent sentences took longer as they seemed inconsistent. These results seem consistent with Ausubel's notion of advance organizers.

## 2.4 Story Grammars

### 2.4.1 Rumelhart

Rumelhart's (1975) generative grammar for stories and set of summarization rules was one of the first text grammars to be developed and tested and has influenced the work of other psychologists (e.g., Mandler & Johnson, 1977; Stein & Glenn, in press; Thorndyke, 1977). In his chapter in Bobrow and Collins (1975) Rumelhart listed two sets of rewrite rules, one called syntactic rules, the other semantic interpretation rules. These are listed in Table 3. The semantic rules are listed using a double arrow. The important differences between the two sets of rules is not obvious. Syntactic rules specify how a sentence is decomposed; the semantic rules are intended to specify the relations among the parts.

Table 3

Rumelhart's Syntactic Rules and [Semantic Interpretation Rules]

(1) Store → Setting + Episode
⇒ [ALLOW (Setting, Episode)]

(2) Setting → (States)*
⇒ [AND (State, state,.....)]

(3) Episode → Event + Reaction
⇒ [INITIATE (Event, Reaction)]

(4) Event → {Episode | Change-of-state | Action | Event + Event}
⇒ [CAUSE ($Event_1$, $Event_2$) or ALLOW ($Event_1$, $Event_2$)]

(5) Reaction → Internal Response + Overt Response
⇒ [MOTIVATE (Interval-response, Overt Response)]

(6) Internal Response → {Emotion | Desire}

(7) Overt Response → {Action | (Attempt)*}
⇒ [THEN ($Attempt_1$, $Attempt_2$,.....)]

(8) Attempt → Plan + Application
⇒ [MOTIVATE (Plan, Application)]

(9) Application → (Preaction)* + Action + Consequence
⇒ [ALLOW (AND (Preaction, Preaction,.....),
{CAUSE | INITIATE | ALLOW} (Action, Consequence))]

(10) Preaction → Subgoal + (Attempt)*
⇒ [MOTIVATE [Subgoal, THEN (Attempt,.....)] ]

(11) Consequence → {Reaction | Event}

Essentially a story is thought to consist of an episode in a particular setting. An episode consists of an event plus a reaction, where an event might consist of a change of state or an action, etc. Rumelhart gave

several examples of stories and fables that seem to fit his grammar. He does not claim that the grammar will fit all stories, especially "more complex multi-protagonist stories." However, he states that it will account for a wide range of simple stories.

Rumelhart's grammar can be applied to simple stories with only limited success. For example, the first rule allows a story to be rewritten into a setting plus an episode. The grammar allows for more than one episode because events can be rewritten into an episode; however, a story is allowed only one setting. Many simple stories have, in fact, different settings for different episodes. More importantly, the tree structure that is generated from the grammar fails to capture the relative importance of various episodes. That is, a second episode must be subordinate to the first. In some stories, however, a second episode can seem at least as important as the first even if the first episode caused the second. Therefore, this second episode should be represented at as high a level in the tree structure.

The representation Rumelhart offers fails to capture critical inferences or interpretations that readers derive from a story. For example, in one fable that he represents (Rumelhart, 1977), a dog loses his meat through greed. The dog saw his reflection in the water and, thinking it was a different dog with meat, he snapped at the second piece of meat, only to lose his own. The reader comprehends the passage by noting something like "greed is counter-productive" or "be satisfied with what you have." Rumelhart's schema does not account for this derived, higher-level notion.

In addition to developing a generative grammar for stories, Rumelhart has been concerned with specifying rules for summarization of stories.

His 1975 paper gives a formal set of summarization rules. When the story is represented appropriately in a tree structure, a summary of the story would include the top nodes of the tree, omitting the lower branches which contain details. A summary of an episode involving the protagonist, P, would include a summary of how P tried to get the goal and a summary of the resulting outcome. If the attempt was successful, the summary might be "P got the goal, G, by a certain method, M."

Recall data and summarization protocols have been examined for sets of stories that fit Rumelhart's grammar. Information high in his tree structures should be recalled more often than information (details) in the lower branches. Rumelhart considers the data of Thorndyke (1977) and Meyer (1975) as support for this basic idea. He has also collected data of his own, looking at subjects' recall to stories after the subjects have summarized the stories. He found that recall was highly correlated with level in the representation hierarchy; a proposition was predicted to be recalled if a proposition lower in the hierarchy was recalled. The conditional probability of a proposition being recalled given that it was predicted to be recalled was about 95%, and the conditional probability of a proposition being predicted given that it was recalled was about 88%. These results seem similar to Meyer's: Information high in his tree structure was recalled as was information high in her content structure; information that is lower in structural importance, and therefore not predicted to be recalled by Rumelhart, was output 12% of the time. Meyer also found some recall of low-level material.

Although Rumelhart has accounted for much of the variance in subjects' recall using his grammar, one must still be cautious in making conclusions on the basis of its predictive power. All the stories he used in his experiment he designed (or selected) to fit the grammar. It is easy to find stories that will not fit the grammar.

2.4.2 Thorndyke

Thorndyke's grammar is adapted from Rumelhart's and bears a great deal of resemblance to it, except he emphasizes the goal structure more than Rumelhart. The main character is seen as trying to solve a problem or obtain a goal. Thorndyke demonstrated the importance of goal structure by rewriting stories that initially fit the grammar to violate various aspects of it. Recall was best in the untouched version of a story for which the goal structure was clearest. It was somewhat worse if the theme or goal was transposed to the end with all the subgoals removed; it was worse yet with no theme; the version that changed the story to consist only of stative and single action sentences had the poorest performance.

Thorndyke showed that recall of a second story was facilitated by first reading and recalling a story with the same structural representation. Repetition of characters in two stories with different representations (of plot) was detrimental. This result is similar to the one reported by Bower (1974), mentioned earlier. The facilitation could have arisen from savings in time needed to process structure, as Thorndyke suggests. This would allow more time to process the content of the story and presumably improve memory for it. On the other hand, the improvement in recall may have been due to better learned retrieval cues. That is, the second recall would be aided by a better

learned story structure to help guide retrieval of the content. It is also possible that both factors were operating. Thorndyke also concluded that content affected recall independently of structural organization. Imagery ratings were correlated with recall of a passage even when factors of structural complexity were partialed out.

### 2.4.3 Mandler and Johnson

Mandler and Johnson (1977) have adapted Rumelhart's grammar and have constructed a story grammar (or "schema" as they call it) that allows more recursion of episodes and is somewhat more complex. They have deleted Rumelhart's semantic structure rules and added a moral at the topic level of the tree structure, similar to van Dijk's analysis.

Mandler and Johnson's approach seems more promising than previous ones in part because of the greater flexibility and recursive capacity of their grammar, the rules for which are given in Table 4. Flexibility is needed in grammars even for "simple" stories, since the class of simple stories has a great deal of variability in structure.

Mandler and Johnson view the story schema as a collection of expectations that are modified as new semantic and syntactic information is processed. The schema is thought to facilitate encoding and retrieval. Because their model attempts to utilize general knowledge as well as knowledge of story grammars, it may be more successful in identifying the constituents of a story. Thorndyke seems to focus on plot structure almost to the exclusion of story content, while Mandler and Johnson analyze which aspects of an episode are most important to the reader. Their predictions concerning recall take content into account.

Table 4

Summary of Rewrite Rules for Mandler & Johnson's Story Grammar

(1) Fable ⟶ Story AND Moral

(2) Story ⟶ Setting AND Event Structure

(3) Setting ⟶ State* (AND Event*)
Event*

(4) State* ⟶ State ((AND State)$^{n}$)

(5) Event* ⟶ Event (( AND / THEN / CAUSE Event)$^{n}$)((AND State)$^{n}$)

(6) Event Structure ⟶ Episode ((THEN Episode)$^{n}$)

(7) Episode ⟶ Beginning CAUSE Development CAUSE Ending

(8) Beginning ⟶ Event*
Episode

(9) Development ⟶ Simple Reaction CAUSE Action
Complex Reaction CAUSE Goal Path

(10) Simple Reaction ⟶ Internal Event ((CAUSE Internal Event)$^{n}$)

(11) Action ⟶ Event

(12) Complex Reaction ⟶ Simple Reaction CAUSE Goal

(13) Goal ⟶ Internal State

(14) Goal Path ⟶ Attempt CAUSE Outcome
Goal Path (CAUSE Goal Path)$^{n}$

(15) Attempt ⟶ Event*

(16) Outcome ⟶ Event*
Episode

(17) Ending ⟶ Event* (AND Emphasis)
Emphasis
Episode

(18) Emphasis ⟶ State

The experimental predictions of Mandler and Johnson go beyond the traditional ones that are concerned with level in the hierarchical representation. For example, they postulate that causally connected episodes will be better recalled than temporally connected episodes. Recall of sentences in the correct order of presentation depends upon the degree of structure present in the passage, such that stories which have a higher degree of structure will be recalled in more accurate temporal order than less structured stories. Elaborations will be poorly recalled. Omissions and violations of the ideal structure will result in additions and distortions in a subject's recall. Their model includes several deletion transformations. The structure of the famous passage "The War of the Ghosts" used by Bartlett (1932) is an example of a structure that is far from ideal. Mandler and Johnson predicted which aspects of this passage would be forgotten and how modifications of the passage will affect recall. Their predictions concerning the types of information that will be retained was supported by a developmental study. The same trends were observed at all ages, although there were some qualitative differences, e.g., young children almost never recall reactions to events.

### 2.4.4 Stein and Glenn

Stein and Glenn's approach bears a great resemblance to that of Mandler and Johnson. It shares many of its virtues. Both grammars contain rules that define the units in a story and the relations among the units. Unlike Rumelhart or Thorndyke, they allow actions as well as states in their settings. One critical difference noted by Black (1978) is that the Stein

and Glenn grammar does not allow lower level units to be re-written into higher level units, while the others do. Like Mandler and Johnson, they have validated their grammars with developmental data as well as adult data. Both groups have shown that organization of information in the subject's recall of a story does not change with age. It is unfortunate that neither group discussed the similarities and differences of their approaches.

An example of the converging experiments Stein and Glenn have conducted to support their grammar are a series recently published by Glenn (1978). In some studies, she would vary the length of stories, keeping structure constant; in others, she would vary the structure of episodes. Amount recalled and number of elaborations generated was affected by length, but recall organization was not. Output protocols were scored according to which aspect of an episode was recalled more often than another. A consistent pattern emerged such that certain informational categories (e.g., internal response of protagonist vs. reaction of protagonist to a consequence) were recalled more frequently than others. This result was taken as evidence for the validity of the grammar which defined these categories. Mandler and Johnson also analyze the pattern of recall on the basis of type of proposition, while Thorndyke notes only level in the hierarchy.

Stein and Glenn claim that their grammar is "inherently" a theory of processing (e.g., Glenn, 1978). Glenn has gone on to argue that because the pattern of recall for a story is consistent with the story grammar's representation of it, "the story grammar, as a model of processing," is supported. Neither assertion is true. No grammar can implicitly contain a process model.

A story grammar only offers rules for generating acceptable stories, provided that one knows what the components in the grammar refer to.

A processing model is necessary to predict behavior using a grammar. However the need for a model does not imply that one exists "inherently." The fact that percent recall varies significantly and consistently according to constituents identified by the grammar is support for the grammar as a descriptive tool. It is not evidence for it as a processing model. A processing model would specify how input is parsed and how the reader determines which constituents are more important, how the information is recalled, etc.

### 2.4.5 Discussion of Story Grammars

The work on story grammars focuses on syntactic structures common to some simple stories. All of the theorists concerned with story grammars and structural representations of text make cogent arguments concerning the importance of such analyses. Mandler and Johnson state that their long range goal is to be able to predict what people will and will not be able to remember from connected discourse. It is also claimed that story grammars are needed to explain prose comprehension. However, an adequate process model has not been offered to explain how these grammars are involved in comprehension. For example, if the grammars proposed are actually used to help parse a passage, how does the reader know which of the many grammars (fables, narratives, complex stories, texts, etc.) is appropriate for a given input? Perhaps multiple grammars are tried in parallel to determine the correct representation. How does the reader know when a passage

is in violation of a given structure (i.e., unacceptable) as opposed to being an exemplar of another grammar? These are certainly not insurmountable problems, but they have not been touched by those arguing for the usefulness of these grammars.

One might wish to argue that part of the usefulness of grammars is to identify the constituents or basic units of comprehension within a story, i.e., those elements that fall at a terminal node in a representation. This would be useful because any process model must identify the constituents it operates on. However, before the terminal propositions can be identified it is necessary to perform a semantic analysis of them, e.g., decide if a terminal proposition can be a "cause." A parsing scheme that uses the meaning or content of the passage to generate expectations about the next words from the input seems more promising. Schank (1972) argued that syntactic parsers were not adequate to account for humans' ability to comprehend sentences. He argued that parsers should dwell on the meanings of the words, not the form classes of the words. Similarly, comprehension models for passages should dwell on the meaning of sentences and not the possible structural representation. Mandler and Johnson admit that thus far they have stressed syntactic aspects of stories, largely ignoring the semantic content. They have incorporated the "content" of a passage into their theory more than most others, but they need to do much more.

The representation of a story in the comprehender's mind must result from some process operating upon the input. Some of the representations discussed thus far (e.g., Meyer) are intended to reflect the input as the

author intended to convey the message. Other grammars discussed are intended to be predictive of the representation of the input in the comprehender's memory. Exactly how the grammar is represented in the comprehender's memory is not clear, nor is it clear how the grammar is used in building a representation.

The next section will focus on process oriented models, on approaches that emphasize world knowledge and on approaches concerned with the inferences made during comprehension. The criterion used to decide which work is included in the present section and which in the following section is not as clear as it should be. For example, some of the research just described seems as concerned with issues of prior knowledge as some of the work to be discussed. Nevertheless, in order to bring attention to the diversity of approaches to the problem of prose processing, a division had to be drawn somewhere.

## 3. USE OF WORLD KNOWLEDGE TO AID COMPREHENSION

### 3.1 Schemata, Frames and Prototypes

The introduction of the term "schema" to replace the term "grammar" is not unique to Mandler and Johnson. Rumelhart (1976) and Rumelhart and Ortony (1977) have more recently viewed Rumelhart's general plot structure or story grammar as just one grammar or "schema" among many that people have available for everyday functioning. Rumelhart and Ortony discuss schemata in terms that go beyond story comprehension. They define a schema as an abstract representation of a generalized concept or situation. There are schemata for pattern perception (e.g., a face schema), for motor skills (e.g., a juggling schema) as well as for comprehension (e.g., a "give" schema). These schemata operate at all levels. A higher level schema will call a subschema to analyze part of the passage. A representation involving schemata would merely have pointers to subschemata. For example, the "stroll" schema would merely point to the "walk" schema. This is a departure from Rumelhart's earlier views in which, for example, all verbs were "unpacked" for comprehension (Rumelhart & Levin, 1975).

Schemata for stories are like grammars in that they allow one to decide what is and what is not a story and what are the constituents of the story. If a story does not fit any of the available schemata, it is not an "acceptable" story. The schema Rumelhart describes is one in which something happens to the protagonist which sets up a goal. The rest of the story is concerned with trying to achieve the goal. A passage is a story fragment when a goal has been set up and the passage does not have a resolution or an outcome.

The idea of using schemata is actually an old one. Rumelhart and Ortony, mention that Kant (1781) as well as Bartlett (1932) put forth related views. More recently, Chafe (1976) and Winograd (1977) have also adopted the term in trying to describe the mechanisms involved in prose comprehension. According to these theorists, a schema or prototype gives the basic structure of a class of its instantiations. No instantiation matches the prototype of schema perfectly. Chafe believes that when a person stores information about an event in memory, the person encodes it with respect to his prototype of similar events, but that the match is not all or none. He argues that information is stored in a "schematic" rather than a "propositional" form because a person does not report the same experience in the same way on two separate occasions, each report being a variation on the schema. This argument is less than convincing for two reasons. First, it is not clear why a propositional representation is incompatible with novelty in expressing the same event; second, my own impression is that the overlap in the retelling of events, jokes, etc. by a given individual is often more striking than any novelty in his exposition. In fact, such retellings often remind me of a prerecorded message being played again.

Winograd (1977) has also adopted the term schemata but relates it to Minsky's (1975) "frames." The distinction he offers between the two concepts is quite subtle. Frames are definitions; they denote a representation with variables or slots to be filled by a particular instance. Certain slots are optional, i.e., they do not have to be given a value for a particular instance. Others are mandatory and assume default values when not specified

in the story. Schemata, on the other hand, are more flexible in that they allow for partial fits in expressing ideas.

Like Rumelhart, Winograd believes that schemata represent concepts which vary in their levels of abstraction; one schema can be embedded within another. Both Winograd and Rumelhart mentioned briefly when presenting papers at the 1976 Carnegie Symposium that schemata include procedural information for recognizing whether something should be classified as an instance of the schema. How this recognition process would work is not yet clear.[6]

While arguing for the virtues of schemata, Winograd also criticized other approaches, viz., the use of scripts and plans by Schank and Abelson (1975), as being far too rigid and not accounting for many problems in comprehension. On the other hand, Schank and Abelson have specified their comprehension procedures to a far greater extent than have those advocating schemata. It is not at all clear how one would write other schemata, even another plot structure, nor how any of its claimed virtues would be implemented.

### 3.2 Necessary Inferences

The notion that drawing inferences is necessary for comprehending a passage has been getting more attention in the last few years (e.g., Charniak, 1972; Clark, 1975; Frederiksen, 1977; Paris, 1975; Rieger, 1974; Schank and Abelson, 1975; Trabasso and Nicholas, 1977). Most of these papers demonstrate the importance of world knowledge and the drawing of inferences by presenting example stories and showing what inferences were

necessary for comprehension. The papers by Clark, Rieger, and Trabasso and Nicholas actually present classifications of necessary inferences. Clark's work will be discussed in detail because he also presents empirical data to support some of his claims.

### 3.2.1 Clark

Clark (1975) and Clark and Haviland (1976) focus their analysis of inferences on those they call "authorized inferences." These inferences are called "authorized inferences" to indicate that they are intended by the speaker, or "backward inferences" to indicate that they are to connect the current sentence with previous ones. Many of the notions Clark and Haviland discuss are based on Grice's (1975) Co-operative Principle. One aspect of this principle concerns finding the intended antecedent that allows comprehension of the current statement. Finding the antecedent sometimes involves "bridging" or the drawing of inferences by the comprehender. Clark (1975) describes a number of types of briding, some of which are described below.

One form of bridging is determining reference. There are many ways in which this can be done. An interesting way is by means of epithet. An example of an epithetic reference is: "I met a man yesterday. The bastard wanted to stop all governmental support of education." Epithets are restricted in use; "the bastard" could not be replaced by "the doctor." Another type of bridging involves inferring information about something previously mentioned. This also may resolve reference. If one reads "I walked into the room. The ceiling was very high," it is easily inferred

that "ceiling" refers to the ceiling of the room since it is a necessary part of the room. However, if one reads, "I walked into the room. The chandeliers sparkled brightly," the idea that the room has chandeliers has been induced; this time the bridge has conveyed new information. Chandeliers are not a standard attribute in a "room frame."

Sometimes the bridge, i.e., the number of inferences involved in making a connection, can be longer than in the above examples. From the set of sentences "John is a Republican. Mary is slightly daft, too," the implications would be: all Republicans are slightly daft; therefore, John is slightly daft. (Some people also infer that Mary is Republican, although that implication is less clear.) Those are both needed as antecedents to the "given information" that someone other than Mary is slightly daft. The number of inferences the listener assumes he should make to bridge the sentences is the minimal number. There are, undoubtedly, a large number of paths of inferences that could connect two sentences. The comprehender must use the rule of finding the shortest path; otherwise, comprehension would take a long time, and there would not be a unique set of antecedents for a sentence that could be intended by the speaker.

There have been a number of experiments which support the notion that the time to comprehend a sentence is affected by the number of inferences that need to be drawn to do so, (Clark & Haviland, 1976; Clark, 1975). Some sentences that subjects read were preceded by a direct antecedent, others by an indirect antecedent. Comprehension of the identical sentence was faster when preceded by the direct antecedent. For example, subjects

were faster to comprehend the second sentence in "Fran took the beer out of the car. The beer was warm," than in "Fran took the picnic supplies out of the car. The beer was warm" or in "Fran likes beer. The beer was warm" (the last being a control for the word "beer").

Clark also makes the important distinction between backward and forward inferences and authorized and unauthorized inferences. Forward inferences are unauthorized (not demanded by the speaker) and are indeterminate in number (i.e., there is no minimal number rule). These are elaborations or embellishments of the input in the text. Backward inferences that are authorized and follow the minimal number rule are the bridges described above. Unauthorized, backward inferences follow from statements such as Nixon's "I am not a crook." In this case, the speaker did not intend the implication that there is reason to consider that he might be a crook.

### 3.2.2 Schank's Model for Drawing Inferences

Clark's paper implies that there are very few and often only one backward inference to be made per sentence. However, Schank's model for sentence comprehension and for paragraph comprehension (Schank, 1972, 1973, 1975a) suggests that many such inferences are made. These backward inferences tend to concern presupposition and enabling conditions. The fact that a large number of inferences need to be made has always seemed to be one of the weakest aspects of Schank's early models. Schank (1972, 1973) asserted in his papers on conceptual dependence (sentence parsing) that during comprehension the verb is decomposed into its semantic primitives. Decomposition includes inferring the instrument of action when the verb

is one that "takes" an instrument. Even in sentences where the instrument is explicitly stated, more implicit instruments may need to be inferred recursively. For example, the conceptual dependency representation of the instrument in "Fred hit Bill with a stick" would be that "Fred did something with a stick," and the instrument for that would be that "Fred grabbed the stick" which implies that "Fred moved his hand toward the stick" and so on.[7]

In a similar vein, Rieger (1974) a student of Schank, developed a scheme in which a seemingly limitless number of "real world inferences" would be made during comprehension. As he states:

> When the current (Conceptual Memory) is turned loose, it will often generate upwards of 100 inferences from a fairly banal stimulus such as "John gave Mary the book." (p. 31)

Rieger described a large number of classes of inferences (16) which include specification, resultative, causative, motivational, enabling, knowledge propagation (inferring what other knowledge an actor must know), and normative inferences. He believes there are even more classes he did not identify and that all these classes generate spontaneous inferences seeking out relevant context. Rieger does give some notions about how the proliferation of inferences will be cut off, but these are not worked out.

Schank (1975a) uses many of these same inference classes in a paper on paragraph comprehension. One motivation for Schank's decision to develop representations for paragraphs was that sentence comprehension involved the

generation of too many inferences; embedding a sentence in a paragraph limits the number of relevant inferences. He feels that in paragraph comprehension, one only needs to generate those inferences that will connect sentences of the story together; one must find enabling conditions, i.e., causal chains that allow an event to occur. He distinguishes between absolutely and reasonably necessary conditions (ANCs and RNCs, respectively). Examples of ANCs for the phrase "John began to mow his lawn" would be John having a lawn, possessing a lawn mower, John being alive, etc. Reasonably necessary conditions would be that it is not cold, nor rainy, nor is there any snow on the ground, etc. None of these conditions need to be stated in the paragraph that includes the phrase. All these statements could be satisfied by "normality assumptions." If the ANCs can be satisfied by "normality assumptions," one only checks to be sure that none of these assumptions are violated by some other assertion in the paragraph.

### 3.3 Scripts, Plans and Goals

Schank and Abelson (1975; 1977; Abelson, 1975) proposed that people utilize situation-specific knowledge, called scripts, in order to make inferences and determine if something seems plausible or reasonable in a particular context. Normality assumptions will vary from context to context (or script to script). The model that uses scripts is implemented in a computer program that summarizes, paraphrases, and translates stories (Schank, 1975b). It can also answer questions about the story concerning what happened, when did something happen and why did it happen.

In addition to scripts, which capture the essence of a stereotyped sequence of actions or events for a well-known situation, there are other comprehension mechanisms such as plans. Plans handle novel situations. Plans describe the set of choices associated with accomplishing a goal. When a plan is frequently implemented, it becomes a script. Most people do not have a "how to become president" script or a "what to do when the house burns down" script but could implement a plan. An example of a script common to most Americans is a restaurant script. Consider (1) below:

(1) John went into a restaurant. He ordered a hamburger and a coke. He asked the waitress for the check and left.

Many aspects of this standard script have been omitted, but are automatically inferred. If the reader of (1) were asked at a delay "Did John eat a hamburger?" the response would probably be "yes," although that was never explicitly stated. One way to tell if an episode is a situational script is by use of the "reference test" (Schank, 1975c). When a definite article precedes a noun not previously mentioned, it is probably acceptable because the referents have been implicitly introduced by the script. Compare (2) and (3):

(2) John went to a restaurant. He asked the waitress to tell the chef to cook him a hot dog.

(3) John went to a restaurant. He struck up a conversation with the bus driver.

In (3) we would feel uncomfortable using "the" with bus driver and might try to augment the script to explain the anomaly.

Scripts are a set of scenes and within each scene there is a causal chain such that each action enables the next. For example, in the restaurant script, asking for the check enables the patron to receive the check. Many parts of a script can be modified and, in fact, frequently are. There are only a few aspects of a script, e.g., ingesting food, which are considered critical. Often there are deviations. Consider scenario (4):

(4) John went to a restaurant. He ordered a hamburger. It was cold when the waitress brought it. He left her a very small tip.

Since there are two deviations from the typical script, viz., "cold" and "very small," the processor would hypothesize that the deviations are related. In general, however, deviations from scripts are handled by "what if" or "whifs" associated with every situational script.

Even though we do not have scripts for every situation encountered in a story, we do need relevant conceptual structure to understand a given input and infer the appropriate connections. Consider the following passage:

(5) John knew his wife's operation would be very expensive. There was always Uncle Harry...he reached for the suburban phone book.

It is unlikely that we have a "paying for an operation" script, yet we might have a "raising money for important expenditures" script. That is, the comprehension of (5) in terms of inferences made would not be very different if "son's education" were substituted for "wife's operation." In each case there is a general goal state. Trying to achieve a goal involves a plan.

A plan is a series of actions that will hopefully realize a goal. Much like the General Problem Solver (Newell & Simon, 1972), plans to achieve the

goal state go through intermediate states, trying to reduce differences between the goal state and current state. Plans consist of a set of "deltacts," actions or subplans. When a set of actions is used often enough together, the set becomes a script. Examples of general purpose subgoals used to achieve a desired state are: D-CONT--a change in the control of an object; D-KNOW--a change in what an actor knows; D-PROX--a change in the proximity relations of objects and actors. Actions that are part of a chain involved in a particular action used for a general purpose belong to a "planbox." Certain plans may invoke a planbox to satisfy a goal. For example, in order to gain control of an object (D-CONT) one might first ask a person for it, bargain for it, threaten a person for it, steal it, or overpower a person in order to take it. Each of these actions involve a number of prerequisites and result in changes of state if the actions succeed. The preconditions that the actor can affect (e.g., one must be "near" the person one is going to ask, or at least physically capable of asking) as well as the resulting state changes are listed in the planbox.

Sometimes in the middle of a story, the actions that a character is pursuing will be frustrated and there will be an abrupt shift in plans. In order to better understand why the actor would shift from the one activity to another, Schank and Abelson described a theory of goal substitution, goal forms, goal initiations, etc. When the system cannot find an appropriate script, it tries to induce what the character's plans would be using the goal monitor. In addition to keeping track of the motivating influence of the character's goals, the monitor must recognize when a goal has been triggered,

makes predictions about what events will be caused by it and keeps track of a goal's fate.

In Schank and Abelson's system, the representation of a story has three levels of description: (1) the lowest level is conceptual dependency (CD), described earlier; (2) main conceptualizations, (MAINCONS), describe the important aspects of a scene and of a script; (3) the highest level is the knowledge structure (KS) which has the script, plan or theme information. The CD level is in some ways similar to Kintsch and van Dijk's micro-structure, while the MAINCONS are like the macro-structure. MAINCONS generate expectations. The three levels are connected by pointers. Like Kintsch and van Dijk and others, Schank and Abelson do not think the lowest level is involved in summarizing. Schank has also predicted that items most likely to be forgotten are those that are at "dead ends" in causal chains. However, unlike Kintsch and van Dijk, Schank has never empirically tested his ideas.

A major weakness of script theory is devising a reasonable scheme for evoking a script -- how does the program know if a script is the correct one? Sometimes several scripts seem appropriate concurrently (e.g., a restaurant script and a romance script). There is no mechanism that allows both scripts to be operating and to make predictions at the same time, let alone interact. These difficulties are not unique to script theory. All extant theories that rely on "bundles" of knowledge to help direct inferences and expectations (e.g., frames and schemata) share the same dilemma. Scripts are currently instantiated when a key phrase has been mentioned (e.g., John went to a restaurant) and a line fitting the script has been mentioned. Most of

the script recognition task has been delegated to the scripts themselves. Others working on simulations have also alluded to procedures attached to schemata that are evoked in the appropriate context so that the relevant schema or script is activated. The workability of these notions remains to be tested in implementation.

### 3.4 The Script-Elaboration Model

The script model described above was described in most detail because it is the processing model with the most promise, in my opinion. The script-elaboration model, a hybrid, borrows heavily from Schank and Abelson's work, as the name implies. The essential difference between this hybrid and the original is the emphasis on the role of elaborative processing. Some of the problems with Schank and Abelson's model discussed above are not solved by these modifications. Nonetheless, the assumption that the reader makes many embellishments during reading has a great deal of psychological plausability and accounts for a large body of data that would otherwise be difficult to explain.

Elaborations are generated that are consistent with the script invoked, but not necessarily true. Consider the following sub-passage.

(6) Bill took Joanne to a steak and ale place, not very fancy but all he could afford. Joanne tried to study the menu intently while Bill gazed at her.

If this were the beginning of a story, i.e., the reader had no prior knowledge about Bill or Joanne, a number of inferences and elaborations would probably occur to the reader. References to "a steak and ale place" and

"menu" would call up a restaurant script and various inferences would be added to the story structure such as Bill and Joanne were shown to a table, given menus by the hostess or waitress, etc. The reader might also imagine the restaurant as being semi-dark, that they were seated in a booth, that the table had placemats rather than a tablecloth, that Bill was infatuated with Joanne, that Joanne was somewhat embarrassed by Bill's infatuation, that this story was a romance story, etc. Another reader might have another set of elaborations entirely, differing in the nature of what was embellished or how it was embellished. Elaborations tend to be highly idiosyncratic, based on one's prior experience with related situations or similar stories.

Elaborations occur at many levels. They can be embellishments of physical descriptions of characters' personalities, imputing goals or intentions to the author, etc. Low level elaborations are generated by accessing from long-term memory salient exemplars of concepts referenced in the story. That is, as each word in a story is read, the corresponding concept in memory may be activated which, in turn, evokes associations to the concept. The instance of a concept stored in memory that would be used to embellish the story will be the most common instance that is consistent with the context described in the story. A "consistent" instance is one that has occurred in context similar to the one described in the story or assumed for the story. The contexts in which exemplars have occurred is also stored in memory.

The generation of elaborations is similar to processes used when scripts are employed to aid comprehension. In both cases, past experience is used

to understand a current situation and in both cases, more information is added to the memory representation of the story. The difference between inferences that represent the omitted main conceptualizations of a script and elaborations is two-fold: Any omitted script action is likely to be inferred, and the default value for the omission (i.e., the inference) is likely to be similar across readers. On the other hand, the number of elaborations generated varies greatly with circumstance and the content of elaborations varies across readers. How much one elaborates depends upon previous experience with the material, inherent interest in the subject matter, understanding of the text, time allotted to read it, concentration and general tendency to elaborate.

Since elaborations are only moderately plausible and are optional, one might ask why people generate elaborations and since they are not necessary, why one should include them in a comprehension model. They serve a number of basic functions, namely, finding connections among sentences, generating expectations about subsequent input, detecting anomalies and aiding retention. Consider the pairs of sentences below:

(7) A West Haven car salesman realized on Thursday that he had not sold many cars all month. The next day, he shot his wife.

(8) A West Haven car salesman realized on Thursday that he had not sold many cars all month. On Friday, a West Haven bank was robbed.

In the first case, elaborations would help establish a "motive," which one might consider a higher level connection. In the second example (8), most

readers would infer that the salesman was the robber, and elaborations are needed to appreciate exactly why he did it. The reader does not typically stop at finding connections between potentially related sentences. He will also speculate on resulting actions, consequences and future problems. In (7) one might expect the salesman to go on trial or go into hiding. In (8) one might expect to hear how he spends the money, how he evades being caught or how he is discovered, etc. In other words, the reader generates expectations concerning subsequent input that severely restricts the original class of possible inputs. Occasionally one is wrong, but for the most part, there are large savings in comprehension time.

Elaborations may also help to detect anomalies. Assertions within a passage may not be in direct contradiction with one another, but the elaborations evoked by the sentences may be inconsistent. Consider (9):

(9) Alice went to Jimmy's house for lunch. Jimmy's mother served them tuna fish sandwiches. Alice liked her sandwich very much and had almost finished it when all of a sudden, her dentures fell out of her mouth.

Although (9) seems anomalous, there is nothing in the passage that is inherently contradictory. Elaborations generated to it, however, might be in direct conflict. For instance, part of my elaborations involved Alice having the smooth skin of a young girl. Then when I imagined dentures falling out, I elaborated a wrinkled face with an old mouth and exposed gums. The anomaly was detected by noting the contradiction of these two facial features. There are of course many other sets of inferences that would also contradict.

Perhaps the most important aspect of the process of elaboration is the benefit to long term retention of the input. Any particular proposition is fragile and may become irretrievable at test. However, if an input proposition had been richly elaborated during reading, then only a few of the redundant propositions need be retrieved at recall to infer the gist of the original input. Suppose, for example, that a reader elaborates passage (6) as follows: "I imagine that the steak and ale place looks a little like Stickney's Restaurant. Perhaps Joanne only pretends to be engrossed by the menu because she is embarrassed by Bill's attentiveness..." etc. Even if the reader could not recall any of the input propositions, he would still have a good chance of recalling the gist of the passage from recalled elaborative fragments. This reconstructive view of memory is not unlike that advocated by Bartlett (1932).

## 3.5 Experimental Evidence in Support of the Elaboration Theory

From the point of view of providing strong empirical support for the elaboration theory, a serious problem is that the amount and direction of elaboration are difficult to manipulate. Nonetheless, there are a number of experiments concerning selectivity in memory for prose which are consistent with an elaboration-plus-reconstruction viewpoint. If subjects have more ability to make certain types of elaborations than others or if subjects are directed to make certain elaborations rather than others, one should see better memory for material consistent with the preferred elaboration and more distortion of material in the direction of the preferred elaboration.

In the classic study conducted by Bartlett (1932), subjects from pre-World War I England studied a northwest coast Indian story, "The War of the Ghosts." Bartlett obtained what he interpreted as systematic distortion of the material in the direction of the knowledge of his subjects. This distortion took the form of additions to the material that made the story more consistent with the world view of his subjects, deletion of inconsistent information, and transformations of inconsistent information to make it more consistent with prevailing beliefs.

There has been a long history of debates (e.g., Gould & Stephenson, 1967; Anderson & Bower, 1973; Spiro, 1975) over the extent to which Bartlett's subjects were really misremembering and the degree to which they were knowingly confabulating in response to perceived task demands. It seems that, at least to some degree, subjects are aware of their distortions and are able to assign lower confidence to these than to veridical recalls. However, this debate misses an important point: The behavior of subjects in Bartlett's task is typical of prose processing. Normally, the reader does not make distinctions between what was actually read in a passage and what is a plausible inference. With most stories the inferences made are plausible extentions of the story and are not distortions. It was Bartlett's clever story selection that served to highlight the elaborative behavior of subjects. The experiments described earlier by Sulin and Dooling (1974) and Dooling and Christiaansen (1977) that showed greater confusion to thematically related foils suggest that distortion can occur as a reconstructive process or as an encoding process.

Bower (1976) reports an interesting experiment looking at the effect of prior information on subjects' memory for a passage. Subjects were given a story that consisted of episodes. Half of the subjects were given prior information that would suggest an unusual interpretation of some of the sub-passages. For instance, subjects might be told that the main character (a college co-ed) had just found out that she was pregnant. The story follows the heroine through five episodes: making a cup of coffee in the morning, visiting a doctor, attending a lecture, going shopping in a grocery store, and attending a cocktail party. The meaning of these episodes can be very different depending on whether or not we view the heroine as pregnant.

Subjects given the interesting interpretation recalled many more inferences appropriate to the pregnancy theme. However, they also recalled more of those episodes related to the theme. This result is what would be expected if subjects had used the information about pregnancy to elaborate. These elaborations should make the text information more redundant and introduce additional inferences.

Hayes (Note 1) has found a similar correlation between number of intruded inferences and overall memory for text. Hayes and his colleagues tried to find out what mechanisms allow some people to remember more than others. They pre-tested subjects on their memory for various historical facts and then classified them as those who remember a lot of history and those that do not. The subjects were then given a fictitious history passage to read. The same subjects who knew more veridical history performed better on a test of the fantasy history passage. Subjects were also asked to free recall the passage that they had read. Not only did the subjects with better

history memory recall more, they also "recalled" many elaborations that were not asserted. These elaborations were not simple paraphrases of the passage, nor were they simple inferences. The subjects classified as having poor memory for history offered almost no elaborations. From this finding, Hayes conjectured that embellishing the input with elaborations promotes better retention.

Schallert (1976) indirectly provides evidence consistent with the notion that elaborations help retention and that the amount of elaboration generated can be influenced by instruction. Subjects were given ambiguous passages that were either biased by prior information or were not biased. She found that subjects in the biased group remembered more information consistent with the bias than those who did not receive prior information. She also introduced a "depth of processing" manipulation in which subjects either processed the sentence at a "shallow" level (counting four letter words in the passage) or at a "deep" level (rating for ambiguity). Biased subjects were more likely to remember consistent information when they were processing the material at the semantic level. It is reasonable to assume that subjects would be generating more elaborations under semantic orienting instructions (see Anderson and Reder, in press), and the hybrid model would claim that elaborations are responsible for the bias found in recall. Therefore, one would expect to find a greater bias in recall for the "deep" processing group. In other words, Schallert's data support the notion of elaborative processing in comprehension because of the interaction of mode of processing and bias.

Brown, Smiley, Day, Townsend, and Lawton (1977) conducted a study with children that supports the elaboration theory by showing that the types of

elaborations generated can be manipulated and that generating more elaborations improves recall. The study indicates that if teachers provide students with background knowledge, students are likely to remember more of the material presented to them. In the Brown et al. study, children in various grades were presented with information about a ficitious tribe called the "Targa" or learned about people from Spain. Those who learned about the Targa were either told that the tribe consisted of Eskimos or desert Indians. A week later all groups read a story about a young boy from the Targa tribe, and no mention was made about what they had studied the preceding week. Of those receiving relevant background information, intrusions and biases in interpretations of ambiguous sentences were consistent with the orientation given earlier. More important, those subjects in the Spanish control recalled significantly less of the veridical material. The usefulness of the background material was evident at all age levels. However, not only did older children recall significantly more veridical information, they recalled significantly more elaborations (had more intrusions) consistent with the background material. In other words, both recall and number of elaborations increased with age.

Owens and Bower (1977) conducted a study that indicates how perspective on a passage can affect memory for the input. Other studies discussed earlier manipulated prior knowledge about concepts in the passage or manipulated the focus of attention by the questions asked during reading. This study was more subtle in that the first few lines of the story caused the reader to identify with one character or the other, depending on which

character was introduced first. Mishaps were described in the story without specifying who was to blame for the accidents. On a subsequent recognition test, subjects were asked to judge which statements were presented in the story. Subjects were much more inclined to false alarm to a statement that imputed blame for the mishap to the character with whom they did not identify, and much less likely to false alarm to statements putting fault on the character with whom they did identify. Similar results have been obtained by Abelson (1976) and Anderson and Pichert (1977). These results are consistent with the notion that readers elaborate material in a fashion that is consistent with their wishes, prejudices or perspective. And there is, of course, some tendency to confuse elaborations with presented input.

In a study of mine (Reder, 1976), I manipulated more directly the amount of elaboration given to prose material. An earlier study of mine indicated that subjects have very good memory for the sentences used in the same experimental stories. So, the dependent measure chosen was the speed with which subjects can make judgments about a story, not the accuracy. The task demanded of subjects was to make plausibility judgments, which seems to resemble everyday tasks more than verbatim memory judgments. Three factors were manipulated orthogonally. With the use of the additive factors logic, I determined that subjects do not search for a specific fact, but rather for a set of propositions that will be relevant for judging plausibility. The speed with which these propositions become available is a function of the number of elaborations generated.

Just as the number of relevant elaborations is postulated to affect retention of the input, amount of relevant elaborations is thought to affect

the speed of retrieval of relevant information. This notion is based on the assumption that searching through memory for relevant information takes time and that the greater the proportion of relevant to irrelevant, information the faster a useful or relevant fact can be found; hence, faster RTs occur with more elaborations. (See Reder, 1976, for a fuller discussion.)

### 3.6 Discussion

The approaches represented in this last section differ from others in a number of respects. First, artificial intelligence models are either implemented on a computer or are purported to be sufficiently specified to be implemented. Second, the models in this section tend to be more concerned with the reader's use of world knowledge than the use of knowledge of language to understand a passage.

Both of these qualities make the models in this last section important advances. When models are implemented, inherent weaknesses in them and issues that were overlooked are quickly brought into focus: A bad program will not run or will produce the "wrong" behavior. In other words, a program serves as a check on the viability of the model that generated it. On the other hand, sometimes the motivation for a particular feature of a model may be computational ease, and the feature may not seem psychologically plausible. Often when computer implementations of models do not work properly, patches are made to the program to make them work. When more and more patches are added to the program, it begins to seem very ad hoc. It also becomes incomprehensible as a model. The modifications to the program tend to have no psychological plausibility whatsoever. Furthermore, AI models are not

always programmed and thus one has neither empirical support nor a computer implementation as confirmation of their viability.

The second advantage of these models, the concern with world knowledge, is unrelated to the issue of whether a model is implemented. My only reservation here concerns the apparent unequal treatment of world knowledge and linguistic knowledge. It seems that some theories are only concerned with the plausibility of the former. The result is that the parser is full of special purpose rules to handle the set of sentences to be processed and they have no relation to linguistic theory nor any generality.

Despite these criticisms, I feel the research in this section is the most commendable. It incorporates world knowledge and is concerned with processing the input, not merely representing it. The most useful theories will be those that predict (1) what information will be best retained from a story, and (2) how to improve a reader's memory for and comprehension of a passage. Approaches concerned with story grammars seem less able to address (2).

The story grammar theories and the processing models share a common flaw: none seem to generalize to many texts or stories. If these theories are to have practical application to education (or even be of more theoretical interest), they must strive to handle a larger class of prose. Hopefully, the complexity of a model will not keep pace with increases in generality.

## 4. GENERAL CONCLUSIONS AND IMPLICATIONS FOR EDUCATION

To some extent, the research discussed in this review has implications for the more applied questions with which we began. For instance, in order to test the facilitation of advance organizers one needs a model of the structure of the text. The thought is that prior exposure to high level concepts will facilitate retention of lower level concepts. There is enough similarity among the grammars described so that one could fairly easily determine which propositions in a text are "high level concepts" and which are "low level concepts."

The processing models, especially the script-elaboration model can explain many of the results concerned with asking questions to improve subsequent performance. From Hayes' data, we learned that there may be inherent differences among readers such that good readers elaborate more than poor readers. However, there is also evidence that the extent to which one elaborates can be manipulated by experimenters (or teachers) and consequently affect performance.[8] Research by Frase, Rothkopf and others, reviewed earlier, demonstrates that asking certain questions at specific times can improve performance. Reder (1976) also found that asking either the same question or a different one that focused attention improved performance substantially in terms of latency to respond. The reaction time advantage for a focusing question over no question increases with delay of final test. The increase in advantage for the focusing questions over delay is to be expected if one assumes that retention of input deteriorates with time.

Schallert's data provides indirect evidence of how different demands on students during study will result in different memory. And Bower's result indicates that teachers should try to make a passage more interesting by relating it to something students already know about. Of course, determining what is "interesting" to a given student is another problem.

The story grammar research also indicates that passages must be "readable" for adequate comprehension and retention. Readability in this context is defined as conformity with the grammar. Work such as Meyer's demonstrates that any point that the author considers important must be integrated in such a way that it is high in the underlying structure that represents the passage. She also showed that merely "flagging" or "signaling" a fact as important is not nearly as effective at insuring its retention.

There is a consensus among researchers studying prose comprehension that much of what is intended by the author is implicit in the passage and that explicit statements in the text are not of equal importance. In order to improve comprehension, then, three major sub-goals must be achieved: (a) train students to automatically infer the implicit information intended that is likely to be necessary for comprehension of subsequent input; (b) teach students to isolate those aspects of the text that seem important and elaborate upon them at the expense of full attention on aspects that seem of less consequence; (c) insure that students have sufficient knowledge of the concepts referred to in the passage so that they can draw required inferences and further elaborate the input. They should learn to recognize when they do not have enough background knowledge and must read something else first.

The first two goals require that a student learn new procedures. The last requirement concerns declarative knowledge rather than procedural. It is not clear which type of knowledge would be easier to convey. Nor is it by any means certain that cognitive science can offer any novel advice on how to improve these skills or transmit the requisite declarative knowledge.

One major difference between good and poor readers is the speed with which they make inferences and elaborations. Assuming that making these inferences is a skill like any other, the most effective way for a person to be faster at the skill is to practice it over and over. I speculate that skills involved in listening to stories are the same skills that determine reading performance in upper grades. That is, most reading problems have nothing to do with initial processing (perception) of the input. Therefore, children could get a lot of practice at the verbal skills of drawing inferences and elaborating if parents would read stories to them frequently. Watching television gives relatively little practice in the verbal skills that will be necessary for reading.

## 4.1 Summing Up

This literature review has stressed the need for more attention to the elaborative aspects of comprehension. The inferences and elaborations that are made during comprehension should be reflected in a structure that is intended to represent how a passage is stored in memory, and a mechanism should be specified for how these inferences are made and why they are made. The hybrid model proposed is only a skeleton of a complete theory of prose comprehension. The exact structure of scripts, how scripts are related in

memory, what other memory structures exist, how representations of stories are stored, etc. all need to be further specified.

Despite the sketchiness of the hybrid model, there are a number of conclusions that can be made that should be of interest to anyone involved in the theory of or application to prose comprehension. A person's comprehension and memory for a passage can be affected by a number of manipulations. The "inherent" differences among readers seems to reflect, in part, differential propensities to elaborate; nevertheless, educators should continue to try to influence the extent of elaboration during reading. Experiments have shown that certain questions asked in the appropriate position can influence the memory for certain facts. Prior knowledge, and inherent interest in the subject matter, orienting instructions, etc. can also affect retention and comprehension.

Reference Note

1. Hayes, J. R. Personal communication, July 1977.

Bibliography

Abelson, R. P. Concepts for representing mundane reality in plans. In D. G. Bobrow & A. M. Collins (Eds.), Representation and understanding: Studies in cognitive science. New York: Academic Press, 1975.

Abelson, R. P. Script processing in attitude formation. In J. S. Carroll & J. W. Payne (Eds.), Cognition and social behavior. Hillsdale, N.J.: Erlbaum, 1976.

Anderson, J. R., & Bower, G. H. Human associative memory. Washington, D.C.: Hemisphere Press, 1973.

Anderson, J. R., & Reder, L. M. An elaborative processing explanation of depth of processing. In L. S. Cermak & F. I. M. Craik (Eds.), Levels of processing in human memory. Hillsdale, N.J.: Erlbaum, in press.

Anderson, R. C. How to construct achievement tests to assess comprehension. Review of Educational Research, 1972, 42, 145-170.

Anderson, R. C., & Biddle, W. B. On asking people questions about what they are reading. In G. H. Bower (Ed.), The psychology of learning and motivation (Vol. 9). New York: Academic Press, 1975.

Anderson, R. C., & Carter, J. F. Retroactive inhibition of meaningfully-learned sentences. American Educational Research Journal, 1972, 9, 443-448.

Anderson, R. C., & Pichert, J. W. Recall of previously unrecallable information following a shift in perspective (Tech. Rep. No. 41). Urbana: University of Illinois, Center for the Study of Reading, April 1977. (ERIC Document Reproduction Service No. ED 142 974)

Ausubel, D. P. The use of advance organizers in learning and retention of meaningful verbal material. Journal of Educational Psychology, 1960, 51, 267-272.

Ausubel, D. P., & Fitzgerald, D. Organizer, general background and antecedent learning variables in sequential verbal learning. Journal of Educational Psychology, 1962, 53, 243-249.

Bartlett, F. C. Remembering: A study in experimental and social psychology. Cambridge, England: Cambridge University Press, 1932.

Bjork, R. A., & Jongeward, R. H. Rehearsal and mere rehearsal. Unpublished manuscript, University of Michigan, 1973.

Black, J. Theories of story memory structure. Unpublished manuscript, Stanford University, 1978.

Bobrow, D. G., & Collins, A. M. (Eds.). Representation and understanding: Studies in cognitive science. New York: Academic Press, 1975.

Bobrow, D. G., & Winograd, T. An overview of KRL, a knowledge representation language. Cognitive Science, 1977, 1, 3-46.

Bobrow, S., & Bower, G. H. Comprehension and recall of sentences. Journal of Experimental Psychology, 1969, 80, 455-461.

Bower, G. H. Selective facilitation and interference in retention of prose. Journal of Educational Psychology, 1974, 66, 1-8.

Bower, G. H. Comprehending and recalling stories. Section 3, Presidential Address presented at the meeting of the American Psychological Association, Washington, D.C., September 1976.

Bransford, J. D., & Franks, J. J. The abstraction of linguistic ideas. Cognitive Psychology, 1971, 2, 331-350.

Bransford, J. D., & Johnson, M. K. Considerations of some problems of comprehension. In W. Chase (Ed.), Visual information processing. New York: Academic Press, 1973.

Bransford, J. D., & McCarrell, N. A sketch of a cognitive approach to comprehension: Some thoughts about understanding what it means to comprehend. In W. B. Weimer & D. S. Palermo (Eds.), Cognition and the symbolic processes. Hillsdale, N.J.: Erlbaum, 1975.

Brown, A. L., Smiley, S. S., Day, J. D., Townsend, M. A. R., & Lawton, S. C. Intrusion of a thematic idea in children's comprehension and retention of stories. Child Development, 1977, 48, 1454-1466.

Bruning, R. H. Effects of review and testlike events within the learning of prose material. Journal of Educational Psychology, 1968, 59, 16-19.

Carroll, J. B. Learning from verbal discourse in educational media: Review of the literature (Final Report, Project 7-1069). Princeton, N.J.: Educational Testing Service, 1971.

Chafe, W. L. Creativity in verbalization and its implications for the nature of stored knowledge. In R. O. Freedle (Ed.), Discourse production and comprehension. Hillsdale, N.J.: Erlbaum, 1976.

Charniak, E. Toward a model of children's story comprehension. MIT Artificial Intelligence Laboratory Mimeo, December 1972.

Clark, H. H. Bridging. In R. C. Schank & B L. Nash-Webber (Eds.), Theoretical issues in natural language processing. Proceedings of conference at the Massachusetts Institute of Technology, June 1975.

Clark, H. H., & Haviland, S. E. Comprehension and the given-new contrast. In R. O. Freedle (Ed.), Discourse production and comprehension. Hillsdale, N.J.: Erlbaum, 1976.

Cofer, C. N. A comparison of logical and verbatim learning of prose passages of different lengths. American Journal of Psychology, 1941, 54, 1-20.

Colby, B. A partial grammar of Eskimo folktales. American Anthologies, 1973, 75, 645-662.

Collins, A. M., & Quillian, M. R. Retrieval time from semantic memory. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 240-247.

Craik, F. I. M., & Lockhart, R. S. Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 671-684.

Crothers, E. Memory structure and the recall of discourse. In R. O. Freedle & J. B. Carroll (Eds.), Language comprehension and the acquisition of knowledge. Washington, D.C.: Winston, 1972.

Crowder, R. G. Principles of learning and memory. Hillsdale, N.J.: Erlbaum, 1976.

Dawes, R. Cognitive distortion. Psychological Reports, 1964, 14, 443-459.

Dawes, R. Memory and the distortion of meaningful written material. British Journal of Psychology, 1966, 57, 77-86.

Deese, J., & Kaufman, R. Serial effects in recall of unorganized and sequentially organized verbal material. Journal of Educational Psychology, 1957, 54, 180-187.

Dooling, D. J., & Christiaansen, R. E. Episodic and semantic aspects of memory for prose. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 428-436.

Dooling, D. J., & Lachman, R. Effects of comprehension on retention of prose. Journal of Experimental Psychology, 1971, 88, 216-222.

Dooling, D. J., & Mullet, R. L. Focus on thematic effects in retention of prose. Journal of Experimental Psychology, 1973, 97, 404-406.

Fillmore, C. J. The case for case. In E. Bach & R. T. Harms (Eds.), Universals in linguistic theory. New York: Holt, Rinehart and Winston, 1968.

Frase, L. T. Learning from prose material: Length of passage, knowledge of results and position of questions. Journal of Educational Psychology, 1967, 58, 266-272.

Frase, L. T. Questions as aids to reading: Some research and a theory. American Educational Research Journal, 1968, 5, 319-332.

Frase, L. T. Paragraph organization of written materials. The influence of conceptual clustering upon level of organization. Journal of Educational Psychology, 1969, 60, 394-401.

Frase, L. T. Influence of sentence order and amount of higher level text processing upon reproductive and productive memory. American Educational Research Journal, 1970, 7, 307-319.

Frase, L. T. Effect of incentive variables and type of adjunct questions. Journal of Educational Psychology, 1971, 62, 371-375.

Frase, L. T. Maintenance and control in the acquisition of knowledge from written materials. In R. O. Freedle & J. B. Carroll (Eds.), Language comprehension and the acquisition of knowledge. Washington, D.C.: Winston, 1972.

Frase, L. T. Prose processing. In G. H. Bower (Ed.), The psychology of learning and motivation (Vol. 9). New York: Academic Press, 1975.

Frase, L. T., & Kammann, R. Effects of search criterion upon anticipated free recall of categorically related words. Memory and Cognition, 1974, 2, 181-184.

Frederiksen, C. H. Effects of task-induced cognitive operations on comprehension and memory processes. In R. O. Freedle & J. B. Carroll (Eds.), Language comprehension and the acquisition of knowledge. Washington, D.C.: Winston, 1972.

Frederiksen, C. H. Acquisition of semantic information from discourse: Effects of repeated exposures. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 158-169. (a)

Frederiksen, C. H. Effects of context-induced processing operations on semantic information acquired from discourse. Cognitive Psychology, 1975, 7, 139-166. (b)

Frederiksen, C. H. Representing logical and semantic structure of knowledge acquired from discourse. Cognitive Psychology, 1975, 7, 371-458. (c)

Frederiksen, C. H. Inference and the structure of children's discourse. Paper presented at the Symposium on the Development of Discourse Processing Skills, Society for Research in Child Development meetings, New Orleans, 1977.

Gagne, R. M., & Wiegand, V. K. Effects of superordinate context on learning and retention of facts. Journal of Educational Psychology, 1970, 61, 406-409.

Glenn, C. G. The role of episodic structure and of story length in children's recall of simple stories. Journal of Verbal Learning and Verbal Behavior, 1978, 17, 229-247.

Gomulicki, B. R. Recall as an abstractive process. Acta Psychologica, 1956, 12, 77-94.

Gould, A., & Stephenson, G. M. Some experiments relating to Bartlett's theory of remembering. British Journal of Psychology, 1967, 58, 39-49.

Greeno, J. G., & Noreen, D. L. Time to read semantically related sentences. Memory and Cognition, 1974, 2, 117-120.

Grice, H. P. Logic and conversation. In P. Cole & J. L. Morgan (Eds.), Syntax and semantics (Vol. 3): Speech acts. New York: Academic Press, 1975.

Grimes, J. The thread of discourse. The Hague: Mouton, 1975.

Hyde, T. S., & Jenkins, J. J. Recall for words as a function of semantic, graphic, and syntactic orienting tasks. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 471-480.

Kant, E. [Critique of pure reason (2nd ed.)] (K. Smith, Ed. and trans.). London: MacMillan, 1963. (Originally published, 1787; 1st ed., 1781.)

King, D. R. W., & Greeno, J. G. Invariance of inference time when information was presented in different linguistic formats. Memory and Cognition, 1974, 2, 233-235.

Kintsch, W. Notes on the structure of semantic memory. In E. Tulving & W. Donaldson (Eds.), Organization of memory. New York: Academic Press, 1972.

Kintsch, W. The representation of meaning in memory. Hillsdale, N.J.: Erlbaum, 1974.

Kintsch, W. On recalling stories. In M. Just & P. Carpenter (Eds.), Cognitive processes in comprehension. Hillsdale, N.J.: Erlbaum, 1977.

Kintsch, W., & Keenan, J. Reading rate and retention as a function of the number of propositions in the base structure of sentences. Cognitive Psychology, 1973, 5, 257-274.

Kintsch, W., McKoon, G., & Keenan, J. Memory for information inferred during reading. In W. Kintsch, (Ed.), The representation of meaning in memory. Hillsdale, N.J.: Erlbaum, 1974.

Kintsch, W., & Monk, D. Storage of complex information in memory: Some implications of the speed with which inferences can be made. Journal of Experimental Psychology, 1972, 94, 25-32.

Kintsch, W., & van Dijk, T. A. [Recalling and summarizing stories.] Langages, 1975, 40, 98-116.

Kirscher, M. C. The effects of presentation order and repetition on free recall of prose. Unpublished master's thesis, Cornell University, 1971.

Lakoff, G. Linguistics and natural logic. Synthese, 1970, 22, 151-271.

Lakoff, G. Structural complexity in fairy tales. The Study of Man, 1972, 1, 128-150.

Mandler, J. M., & Johnson, N. S. Remembrance of this parsed: Story structure and recall. Cognitive Psychology, 1977, 9, 111-151.

McGraw, B., & Grotelueschen, A. Direction of the effect of questions in prose material. Journal of Educational Psychology, 1972, 63, 586-588.

McKoon, G. Organization of information in text memory. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 247-260.

Meyer, B. J. F. The organization of prose and its effect on recall. Amsterdam: North-Holland Publishers, 1975.

Meyer, B. J. F., & McConkie, G. W. What is recalled after hearing a passage? Journal of Educational Psychology, 1973, 65, 109-117.

Minsky, M. A framework for representing knowledge. In P. H. Winston (Ed.), The psychology of computer vision. New York: McGraw-Hill, 1975.

Montague, W. E., Adams, J. A., & Kiess, H. O. Forgetting and natural language mediation. Journal of Experimental Psychology, 1966, 72, 829-833.

Montague, W. E., & Carter, J. Vividness and imagery in recalling connected discourse. Journal of Educational Psychology, 1973, 64, 72-75.

Myers, J. L., Pezdek, K., & Coulson, D. Effect of prose organization upon free recall. Journal of Educational Psychology, 1973, 65, 313-320.

Myrow, D. L., & Anderson, R. E. Retroactive inhibition of prose as a function of the type of test. Journal of Educational Psychology, 1972, 63, 303-308.

Newell, A., & Simon, H. Human problem solving. Englewood Cliffs, N.J.: Prentice Hall, 1972.

Owens, J. E., & Bower, G. H. Character point of view in text comprehension and memory. Paper presented at the meeting of the American Psychological Association, San Francisco, August 1977.

Paris, S. G. Integration and inference in children's comprehension and memory. In F. Restle, R. M. Shiffrin, N. J. Castellan, H. R. Lindman & D. B. Pisoni (Eds.), Cognitive theory (Vol. 1). Hillsdale, N.J.: Erlbaum, 1975.

Paris, S. G., & Lindauer, B. K. Constructive aspects of children's comprehension and memory. In R. V. Kail & J. W. Hagen (Eds.), Perspective on the development of memory and cognition. Hillsdale, N.J.: Erlbaum, in press.

Paul, I. H. Studies in remembering. Psychological Issues, Monograph 2, Vol. 1, 2, 1959.

Perlmutter, J., & Royer, J. M. Organization of prose materials: Stimulus, storage and retrieval. Canadian Journal of Psychology, 1973, 27, 200-209.

Postman, L. Extra-experimental interference and the retention of words. Journal of Experimental Psychology, 1961, 61, 97-110.

Propp, V. Morphology of the folktale. International Journal of American Linguistics, 1958, 24, Part 3.

Prytulak, L. G. Natural language mediation. Cognitive Psychology, 1971, 2, 1-56.

Reder, L. M. The role of elaborations in the processing of prose. Unpublished doctoral dissertation, University of Michigan, 1976.

Richardson, P., & Voss, J. F. Verbal context and recall of meaningful material. Journal of Experimental Psychology, 1960, 60, 417-418.

Rieger, C. J. Conceptual memory: A theory and computer program for processing the meaning content of natural language utterances. Unpublished doctoral dissertation, Stanford University, 1974.

Roediger, H. L., III. Inhibition in recall from cueing with recall targets. Journal of Verbal Learning and Verbal Behavior, 1973 12, 644-657.

Roediger, H. L., III. Inhibiting effects of recall. Memory and Cognition, 1974, 2, 261-269.

Rothkopf, E. Z. Some theoretical and experimental approaches to problems in written instruction. In J. D. Krumboltz (Ed.), Learning and the educational process. Chicago: Rand McNally, 1965.

Rothkopf, E. Z. Learning from written instructive materials: An exploration of the control of inspection behavior by test-like events. American Educational Research Journal, 1966, 3, 241-249.

Rothkopf, E. Z. Structural text features and the control of processes in learning from written materials. In R. O. Freedle & J. B. Carroll (Eds.), Language comprehension and the acquisition of knowledge. Washington, D.C.: Winston, 1972.

Rothkopf, E. Z., & Bisbicos, E. Selective facilitative effects of interspersed questions on language from written material. Journal of Educational Psychology, 1967, 58, 56-61.

Rothkopf, E. Z., & Kaplan, R. Exploration of the effect of density and specificity of instructional objectives on learning from text. Journal of Educational Psychology, 1972, 63, 309-313.

Rumelhart, D. E. Notes on a schema for stories. In D. G. Bobrow & A. M. Collins (Eds.), Representation and understanding: Studies in cognitive science. New York: Academic Press, 1975.

Rumelhart, D. E. Comprehension and the structure of brief stories. Carnegie Symposium on Cognitive Processes in Comprehension, May 1976.

Rumelhart, D. E. Understanding and summarizing brief stories. In D. La Berge & J. Samuels (Eds.), Basic processes in reading: Perception and comprehension. Hillsdale, N.J.: Erlbaum, 1977.

Rumelhart, D. E., & Levin, J. A. A language comprehension system. In D. A. Norman, D. E. Rumelhart & the LNR Research Group, Explorations in cognition. San Francisco: Freeman, 1975.

Rumelhart, D. E., Lindsay, P. H., & Norman, D. A. A process model for long-term memory. In E. Tulving & W. Donaldson (Eds.), Organization of memory. New York: Academic Press, 1972.

Rumelhart, D. E., & Ortony, A. The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro & W. E. Montague (Eds.), Schooling and the acquisition of knowledge. Hillsdale, N.J.: Erlbaum, 1977.

Schallert, D. L. Improving memory for prose: The relationship between depth of processing and context. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 621-632.

Schank, R. C. Conceptual dependency: A theory of natural language understanding. Cognitive Psychology, 1972, 3, 552-631.

Schank, R. C. Identification of conceptualizations underlying natural language. In R. C. Schank & K. M. Colby (Eds.), Computer models of thought and language. San Francisco: Freeman, 1973.

Schank, R. C. The structure of episodes in memory. In D. G. Bobrow & A. M. Collins (Eds.), Representation and understanding: Studies in cognitive science. New York: Academic Press, 1975. (a)

Schank, R. C. (& the Yale AI Project). SAM--A story understander. Research Research Report No. 3, Yale University, August 1975. (b)

Schank, R. C. The role of memory in language processing. In C. Cofer & R. Atkinson (Eds.), The nature of human memory. San Francisco: Freeman, 1975. (c)

Schank, R. C., & Abelson, R. P. Scripts, plans, and knowledge. Paper presented at the Fourth International Joint Conference on Artificial Intelligence, Tbilisi, USSR, 1975.

Schank, R. C., & Abelson, R. P. Scripts, plans, goals, and understanding: An inquiry into human knowledge structures. Hillsdale, N.J.: Erlbaum, 1977.

Smiley, S. S., Oakley, D. D., Worthen, D., Campione, J. C., & Brown, A. L. Recall of thematically relevant material by adolescent good and poor readers as a function of written versus oral presentation. Educational Psychology, 1977, 69, 381-387.

Spiro, R. J. Inferential reconstruction in memory for connected discourse (Tech. Rep. No. 2). Urbana: University of Illinois, Laboratory for Cognitive Studies in Education, October 1975. (ERIC Document Reproduction Service No. ED 136 187)

Stein, N. L., & Glenn, C. G. An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), Discourse processing: New directions. Norwood, N.J.: Ablex, in press.

Sulin, R. A., & Dooling, D. J. Intrusions of a thematic idea in retention of prose. Journal of Experimental Psychology, 1974, 103, 255-262.

Thomas, E. L., & Robinson, H. A. Improving reading in every class: A sourcebook for teachers. Boston: Allyn and Bacon, 1972.

Thorndyke, P. W. Cognitive structure in human story comprehension and memory. Unpublished doctoral dissertation, Stanford University, 1975.

Thorndyke, P. W. Cognitive structures in comprehension and memory of narrative discourse. Cognitive Psychology, 1977, 9, 77-110.

Trabasso, T., & Nicholas, D. W. Memory and inferences in the comprehension of narratives. Paper presented at a conference entitled "Structure and Process, Models in the Study of Dimensionality of Children's Judgments," Kassel, Germany, June 1977.

Tulving, E. Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), Organization of memory. New York: Academic Press, 1972.

Tulving, E., & Thomson, D. M. Encoding specificity and retrieval processes in episodic memory. Psychological Review, 1973, 80, 352-373.

Walker, B. S. Effects of inserted questions on retroactive inhibition in meaningful verbal learning. Journal of Educational Psychology, 1974, 66, 486-490.

Watts, G. H., & Anderson, R. C. Effects of three types of inserted questions on learning from prose. Journal of Educational Psychology, 1971, 62, 387-394.

Winograd, T. Discourse. In M. Just & P. Carpenter (Eds.), Cognitive processes in comprehension. Hillsdale, N.J.: Erlbaum, 1977.

van Dijk, T. Recalling and summarizing complex discourse. Unpublished Mimeo, University of Amsterdam, March 1975.

van Dijk, T. Macro-structures and cognition. In M. Just & P. Carpenter (Eds.), Cognitive processes in comprehension. Hillsdale, N.J.: Erlbaum, 1977.

van Dijk, T., & Kintsch, W. Cognitive psychology and discourse. In W. U. Dressler (Ed.), Trends in text linguistics. New York: DeGruyter, 1976.

## Footnotes

This paper was prepared while the author was an NSF graduate fellow at the University of Michigan and an NIMH post-doctoral fellow at Yale University. The manuscript has benefited from the comments of J. Anderson, J. Baker, S. Chipman, N. Stein, and R. Tourangeau. The author's address is Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213.

[1]There have been a few notable exceptions. For example, the psychologist Cofer (1941), demonstrated that it is easier to recall the gist of a long passage than to recall it verbatim; and Gomulicki (1956) found that subjects are more likely to free recall the agent and action phrases than the less important, modifying segments.

[2]There is an interpretation of these remarks with which I do not wish to be associated. A common position of late is that the "correct" domain to work in is prose processing or, even more specifically, "story" processing; that working with more "impoverished" material leads to distortions of the normal processing. This point of view and the methodological imperialism it leads to is nonsense. The human engages in a wide variety of behavior. There is no reason to believe that story comprehension of the kind currently in vogue is more representative of normal processes than is free recall. In fact, there are probably more adult behaviors (e.g., remembering a grocery list) that come close to a free recall experiment than adult behaviors that come close to reading one of the very simple stories typically employed. It seems silly to regard any paradigm as prototypical. A theorist may miss

important generalities in human behavior if the tasks studied are restricted to a narrow paradigm.

[3]Letter identification, word identification, phonetic coding, etc. are important aspects of reading, but how they are performed is not addressed in this review. It has been shown (Smiley, Oakley, Worthen, Campione & Brown, 1977; Anderson & Biddle, 1975) that the differences between good readers and poor readers maintain regardless of the input modality of the message to be retained--visual or verbal. Many people's problems with reading are not ones of encoding but of comprehension which they would have whether they listened to the message or read it.

[4]This improvement, however, may be due to guessing since they used a cued recall procedure that required subjects to fill in the missing words of statements. Since the subject easily remembers the recently presented "topic," the additional recall cue may allow him to infer the original fact. A control condition would have been useful in which subjects who never saw the original sentences tried to "recall" to the "retrieval cues."

[5]Although the conclusion offered by Kintsch et al. may be correct, the statistics performed on the data do not warrant this conclusion. They have merely failed to reject the null hypothesis at the two longest delays. The appropriate test would have been to assume, as the null hypothesis, no difference between explicit and implicit inferences over time, i.e., assume the functions to be parallel across delay. If they found an interaction between delay and explicit vs. implicit, they would be more justified in their conclusion.

[6]Bobrow and Winograd (1977) have developed a language called KRL (Knowledge Representation Language) that is supposed to recognize instances of a schema. The details of how this is accomplished have not been given, nor are the limitations of the program known.

[7]At a Sloan sponsored workshop at Yale in June, 1978, Schank indicated that his position on decomposition has been modified. He now believes there exist "primitives" at many different levels. One need not unpack a verb down to the original "primitive acts" postulated by Schank. I find the notion of different levels of primitives somewhat disconcerting. Perhaps primitives now mean "concepts."

[8]One should not conclude from these remarks that teachers have not already been encouraging behaviors that seem consistent with these findings (Frase, 1975). The PQ4R method (Thomas & Robinson, 1972) is used by many reading specialists. This method encourages the reader to a) preview, b) make up salient questions, c) read the text, d) reflect upon it, e) recite it, and f) review it. Previewing is intended to allow the student to better allocate his processing time when the material is read more carefully. Asking questions about the text prior to reading is intended to help the subject better attend to important aspects of the material. The data reviewed here, however, indicates that time is better spent "reflecting" upon the material, that is, answering questions afterwards. Much more work needs to be done in the field of reading research before more specific proposals can replace or clarify this method.

CENTER FOR THE STUDY OF READING

READING EDUCATION REPORTS

No. 1: Durkin, D. Comprehension Instruction--Where Are You?, October 1977. (ERIC Document Reproduction Service No. ED 146 566, 14p., HC-$1.67, MF-$.83)

No. 2: Asher, S. R. Sex Differences in Reading Achievement, October 1977. (ERIC Document Reproduction Service No. ED 146 567, 30p., HC-$2.06, MF-$.83)

No. 3: Adams, M. J., Anderson, R. C., & Durkin, D. Beginning Reading: Theory and Practice, November 1977.

No. 4: Jenkins, J. R., & Pany, D. Teaching Reading Comprehension in the Middle Grades, January 1978.

No. 5: Bruce, B. What Makes a Good Story?, June 1978.

No. 6: Anderson, T. H. Another Look at the Self-Questioning Study Technique, September 1978.

CENTER FOR THE STUDY OF READING

TECHNICAL REPORTS

* Available only through ERIC

*No. 1: Halff, H. M. Graphical Evaluation of Hierarchical Clustering Schemes, October 1975. (ERIC Document Reproduction Service No. ED 134 926, 11p., HC-$1.67, MF-$.83)

*No. 2: Spiro, R. J. Inferential Reconstruction in Memory for Connected Discourse, October 1975. (ERIC Document Reproduction Service No. ED 136 187, 81p., HC-$4.67, MF-$.83)

*No. 3: Goetz, E. T. Sentences in Lists and in Connected Discourse, November 1975. (ERIC Document Reproduction Service No. ED 134 927, 75p., HC-$3.50, MF-$.83)

*No. 4: Alessi, S. M., Anderson, T. H., & Biddle, W. B. Hardware and Software Considerations in Computer Based Course Management, November 1975. (ERIC Document Reproduction Service No. ED 134 928, 21p., HC-$1.67, MF-$.83)

*No. 5: Schallert, D. L. Improving Memory for Prose: The Relationship Between Depth of Processing and Context, November 1975. (ERIC Document Reproduction Service No. ED 134 929, 37p., HC-$2.06, MF-$.83)

*No. 6: Anderson, R. C., Goetz, E. T., Pichert, J. W., & Halff, H. M. Two Faces of the Conceptual Peg Hypothesis, January 1976. (ERIC Document Reproduction Service No. ED 134 930, 29p., HC-$2.06, MF-$.83)

*No. 7: Ortony, A. Names, Descriptions, and Pragmatics, February 1976. (ERIC Document Reproduction Service No. ED 134 931, 25p., HC-$1.67, MF-$.83)

No. 8: Mason, J. M. Questioning the Notion of Independent Processing Stages in Reading, February 1976. (Journal of Educational Psychology, 1977, 69, 288-297)

*No. 9: Siegel, M. A. Teacher Behaviors and Curriculum Packages: Implications for Research and Teacher Education, April 1976. (ERIC Document Reproduction Service No. ED 134 932, 42p., HC-$2.06, MF-$.83)

*No. 10: Anderson, R. C., Pichert, J. W., Goetz, E. T., Schallert, D. L., Stevens, K. V., & Trollip, S. R. Instantiation of General Terms, March 1976. (ERIC Document Reproduction Service No. ED 134 933, 30p., HC-$2.06, MF-$.83)

*No. 11: Armbruster, B. B. Learning Principles from Prose: A Cognitive Approach Based on Schema Theory, July 1976. (ERIC Document Reproduction Service No. ED 134 934, 48p., HC-$2.06, MF-$.83)

*No. 12: Anderson, R. C., Reynolds, R. E., Schallert, D. L., & Goetz, E. T. Frameworks for Comprehending Discourse, July 1976. (ERIC Document Reproduction Service No. ED 134 935, 33p., HC-$2.06, MF-$.83)

No. 13: Rubin, A. D., Bruce, B. C., & Brown, J. S. A Process-Oriented Language for Describing Aspects of Reading Comprehension, November 1976. (ERIC Document Reproduction Service No. ED 136 188, 41p., HC-$2.06, MF-$.83)

No. 14: Pichert, J. W., & Anderson, R. C. Taking Different Perspectives on a Story, November 1976. (ERIC Document Reproduction Service No. ED 134 936, 30p., HC-$2.06, MF-$.83)

No. 15: Schwartz, R. M. Strategic Processes in Beginning Reading, November 1976. (ERIC Document Reproduction Service No. ED 134 937, 19p., HC-$1.67, MF-$.83)

No. 16: Jenkins, J. R., & Pany, D. Curriculum Biases in Reading Achievement Tests, November 1976. (ERIC Document Reproduction Service No. ED 134 938, 24p., HC-$1.67, MF-$.83)

No. 17: Asher, S. R., Hymel, S., & Wigfield, A. Children's Comprehension of High- and Low-Interest Material and a Comparison of Two Cloze Scoring Methods, November 1976. (ERIC Document Reproduction Service No. ED 134 939, 32p., HC-$2.06, MF-$.83)

No. 18: Brown, A. L., Smiley, S. S., Day, J. D., Townsend, M. A. R., & Lawton, S. C. Intrusion of a Thematic Idea in Children's Comprehension and Retention of Stories, December 1976. (ERIC Document Reproduction Service No. ED 136 189, 39p., HC-$2.06, MF-$.83)

No. 19: Kleiman, G. M. The Prelinguistic Cognitive Basis of Children's Communicative Intentions, February 1977. (ERIC Document Reproduction Service No. ED 134 940, 51p., HC-$3.50, MF-$.83)

No. 20: Kleiman, G. M. The Effect of Previous Context on Reading Individual Words, February 1977. (ERIC Document Reproduction Service No. ED 134 941, 76p., HC-$4.67, MF-$.83)

No. 21: Kane, J. H., & Anderson, R. C. Depth of Processing and Interference Effects in the Learning and Remembering of Sentences, February 1977. (ERIC Document Reproduction Service No. ED 134 942, 29p., HC-$2.06, MF-$.83)

No. 22: Brown, A. L., & Campione, J. C. Memory Strategies in Learning: Training Children to Study Strategically, March 1977. (ERIC Document Reproduction Service No. ED 136 234, 54p., HC-$3.50, MF-$.83)

No. 23: Smiley, S. S., Oakley, D. D., Worthen, D., Campione, J. C., & Brown, A. L. Recall of Thematically Relevant Material by Adolescent Good and Poor Readers as a Function of Written Versus Oral Presentation, March 1977. (ERIC Document Reproduction Service No. ED 136 235, 23p., HC-$1.67, MF-$.83)

No. 24: Anderson, R. C., Spiro, R. J., & Anderson, M. C. Schemata as Scaffolding for the Representation of Information in Connected Discourse, March 1977. (ERIC Document Reproduction Service No. ED 136 236, 18p., HC-$1.67, MF-$.83)

No. 25: Pany, D., & Jenkins, J. R. Learning Word Meanings: A Comparison of Instructional Procedures and Effects on Measures of Reading Comprehension with Learning Disabled Students, March 1977. (ERIC Document Reproduction Service No. ED 136 237, 34p., HC-$2.06, MF-$.83)

No. 26: Armbruster, B. B., Stevens, R. J., & Rosenshine, B. Analyzing Content Coverage and Emphasis: A Study of Three Curricula and Two Tests, March 1977. (ERIC Document Reproduction Service No. ED 136 238, 22p., HC-$1.67, MF-$.83)

No. 27: Ortony, A., Reynolds, R. E., & Arter, J. A. Metaphor: Theoretical and Empirical Research, March 1977. ERIC Document Reproduction Service No. ED 137 752, 63p., HC-$3.50, MF-$.83)

No. 28: Ortony, A. Remembering and Understanding Jabberwocky and Small-Talk, March 1977. (ERIC Document Reproduction Service No. ED 137 753, 36p., HC-$2.06, MF-$.83)

No. 29: Schallert, D. L., Kleiman, G. M., & Rubin, A. D. Analysis of Differences Between Oral and Written Language, April 1977. (ERIC Document Reproduction Service No. ED 144 038, 33p., HC-$2.06, MF-$.83)

No. 30: Goetz, E. T., & Osborn, J. Procedures for Sampling Texts and Tasks in Kindergarten through Eighth Grade, April 1977. (ERIC Document Reproduction Service No. ED 146 565, 80p., HC-$4.67, MF-$.83)

No. 31: Nash-Webber, B. Anaphora: A Cross-Disciplinary Survey, April 1977. (ERIC Document Reproduction Service No. ED 144 039, 43p., HC-$2.06, MF-$.83)

No. 32: Adams, M. J., & Collins, A. A Schema-Theoretic View of Reading Comprehension, April 1977. (ERIC Document Reproduction Service No. ED 146 565, 80p., HC-$4.67, MF-$.83)

No. 33: Huggins, A. W. F. Syntactic Aspects of Reading Comprehension, April 1977. (ERIC Document Reproduction Service No. ED 142 972, 68p., HC-$3.50, MF-$.83)

No. 34: Bruce, B. C. Plans and Social Actions, April 1977. (ERIC Document Reproduction Service No. ED 149 328, 45p., HC-$2.06, MF-$.83)

No. 35: Rubin, A. D. Comprehension Processes in Oral and Written Language, April 1977. (ERIC Document Reproduction Service No. ED 150 550, 61p., HC-$3.50, MF-$.83)

No. 36: Nash-Webber, B., & Reiter, R. Anaphora and Logical Form: On Formal Meaning Representations for Natural Language, April 1977. (ERIC Document Reproduction Service No. ED 142 973, 42p., HC-$2.06, MF-$.83)

No. 37: Adams, M. J. Failures to Comprehend and Levels of Processing in Reading, April 1977.

No. 38: Woods, W. A. Multiple Theory Formation in High-Level Perception, April 1977. (ERIC Document Reproduction Service No. ED 144 020, 58p., HC-$3.50, MF-$.83)

No. 40: Collins, A., Brown, J. S., & Larkin, K. M. Inference in Text Understanding, December 1977. (ERIC Document Reproduction Service No. ED 150 547, 48p., HC-$2.06, MF-$.83)

No. 41: Anderson, R. C., & Pichert, J. W. Recall of Previously Unrecallable Information Following a Shift in Perspective, April 1977. (ERIC Document Reproduction Service No. ED 142 974, 37p., HC-$2.06, MF-$.83)

No. 42: Mason, J., Osborn, J., & Rosenshine, B. A Consideration of Skill Hierarchy Approaches to the Teaching of Reading, December 1977. (ERIC Document Reproduction Service No. ED 150 549, 176p., HC-$10.03, MF-$.83)

No. 43: Collins, A., Brown, A. L., Morgan, J. L., & Brewer, W. F. The Analysis of Reading Tasks and Texts, April 1977.

No. 44: McClure, E. Aspects of Code-Switching in the Discourse of Bilingual Mexican-American Children, April 1977. (ERIC Document Reproduction Service No. ED 142 975, 39p., HC-$2.06, MF-$.83)

No. 45: Schwartz, R. M. Relation of Context Utilization and Orthographic Automaticity in Word Identification, May 1977.

No. 46: Anderson, R. C., Stevens, K. C., Shifrin, Z., & Osborn, J. Instantiation of Word Meanings in Children, May 1977. (ERIC Document Reproduction Service No. ED 142 976, 22p., HC-$1.67, MF-$.83)

No. 47: Brown, A. L. Knowing When, Where, and How to Remember: A Problem of Metacognition, June 1977. (ERIC Document Reproduction Service No. ED 146 562, 152p., HC-$8.69, MF-$.83)

No. 48: Brown, A. L., & DeLoache, J. S. Skills, Plans, and Self-Regulation, July 1977. (ERIC Document Reproduction Service No. ED 144 040, 66p., HC-$3.50, MF-$.83)

No. 49: Goetz, E. T. Inferences in the Comprehension of and Memory for Text, July 1977. (ERIC Document Reproduction Service No. ED 150 548, 97p., HC-$4.67, MF-$.83)

No. 50: Anderson, R. C. Schema-Directed Processes in Language Comprehension, July 1977. (ERIC Document Reproduction Service No. ED 142 977, 33p., HC-$2.06, MF-$.83)

No. 51: Brown, A. L. Theories of Memory and the Problems of Development: Activity, Growth, and Knowledge, July 1977. (ERIC Document Reproduction Service No. ED 144 041, 59p., HC-$3.50, MF-$.83)

No. 52: Morgan, J. L. Two Types of Convention in Indirect Speech Acts, July 1977.

No. 53: Brown, A. L., Smiley, S. S., & Lawton, S. C. The Effects of Experience on the Selection of Suitable Retrieval Cues for Studying from Prose Passages, July 1977. (ERIC Document Reproduction Service No. ED 144 042, 30p., HC-$2.06, MF-$.83)

No. 54: Fleisher, L. S., & Jenkins, J. R. Effects of Contextualized and Decontextualized Practice Conditions on Word Recognition, July 1977. (ERIC Document Reproduction Service No. ED 144 043, 37p., HC-$2.06, MF-$.83)

No. 55: Jenkins, J. R., & Larson, K. Evaluating Error Correction Procedures for Oral Reading, June 1978.

No. 56: Anderson, T. H., Standiford, S. N., & Alessi, S. M. Computer Assisted Problem Solving in an Introductory Statistics Course, August 1977. (ERIC Document Reproduction Service No. ED 146 563, 26p., HC-$2.06, MF-$.83)

No. 57: Barnitz, J. Interrelationship of Orthography and Phonological Structure in Learning to Read, August 1977. (ERIC Document Reproduction Service No. ED 150 546, 62p., HC-$3.50, MF-$.83)

No. 58: Mason, J. M. The Role of Strategy in Reading in the Mentally Retarded, September 1977.

No. 59: Mason, J. M. Reading Readiness: A Definition and Skills Hierarchy from Preschoolers' Developing Conceptions of Print, September 1977.

No. 60: Spiro, R. J., & Esposito, J. J. Superficial Processing of Explicit Inferences in Text, December 1977. (ERIC Document Reproduction Service No. ED 150 545, 27p., HC-$2.06, MF-$.83)

No. 65: Brewer, W. F. Memory for the Pragmatic Implications of Sentences, October 1977. (ERIC Document Reproduction Service No. ED 146 564, 27p., HC-$2.06, MF-$.83)

No. 66: Brown, A. L., & Smiley, S. S. The Development of Strategies for Studying Prose Passages, October 1977.

No. 68: Stein, N. L., & Nezworski, T. The Effects of Organization and Instructional Set on Story Memory, January 1978. (ERIC Document Reproduction Service No. ED 149 327, 41p., HC-$2.06, MF-$.83)

No. 69: Stein, N. L. How Children Understand Stories: A Developmental Analysis, March 1978.

No. 76: Thieman, T. J., & Brown, A. L. The Effects of Semantic and Formal Similarity on Recognition Memory for Sentences in Children, November 1977. (ERIC Document Reproduction Service No. ED 150 551, 26p., HC-$2.06, MF-$.83)

No. 77: Nash-Webber, B. L. Inference in an Approach to Discourse Anaphora, January 1978. (ERIC Document Reproduction Service No. ED 150 552, 30p., HC-$2.06, MF-$.83)

No. 78: Gentner, D. On Relational Meaning: The Acquisition of Verb Meaning, December 1977. (ERIC Document Reproduction Service No. ED 149 325, 46p., HC-$2.06, MF-$.83)

No. 79: Royer, J. M. Theories of Learning Transfer, January 1978. (ERIC Document Reproduction Service No. ED 149 326, 55p., HC-$3.50, MF-$.83)

No. 80: Arter, J. A., & Jenkins, J. R. Differential Diagnosis-Prescriptive Teaching: A Critical Appraisal, January 1978. (ERIC Document Reproduction Service No. ED 150 578, 104p., HC-$6.01, MF-$.83)

No. 81: Shoben, E. J. Choosing a Model of Sentence Picture Comparisons: A Reply to Catlin and Jones, February 1978. (ERIC Document Reproduction Service No. ED 150 577, 30p., HC-$2.06, MF-$.83)

No. 82: Steffensen, M. S. Bereiter and Engelmann Reconsidered: The Evidence from Children Acquiring Black English Vernacular, March 1978.

No. 83: Reynolds, R. E., Standiford, S. N., & Anderson, R. C. Distribution of Reading Time when Questions are Asked about a Restricted Category of Text Information, April 1978.

No. 84: Baker, L. Processing Temporal Relationships in Simple Stories: Effects of Input Sequence, April 1978.

No. 85: Mason, J. M., Knisely, E., & Kendall, J. Effects of Polysemous Words on Sentence Comprehension, May 1978.

No. 86: Anderson, T. H., Wardrop, J. L., Hively, W., Muller, K. E., Anderson, R. I., Hastings, C. N., & Frederiksen, J. Development and Trial of a Model for Developing Domain Referenced Tests of Reading Comprehension, May 1978.

No. 87: André, M. E. D. A., & Anderson, T. H. The Development and Evaluation of a Self-Questioning Study Technique, June 1978.

No. 88: Bruce, B., & Newman, D. Interacting Plans, June 1978.

No. 89: Bruce, B., Collins, A., Rubin, A. D., & Gentner, D. A Cognitive Science Approach to Writing, June 1978.

No. 90: Asher, S. T. Referential Communication, June 1978.

No. 91: Royer, J. M., & Cunningham, D. J. On the Theory and Measurement of Reading Comprehension, June 1978.

No. 92: Mason, J. M., & Kendall, J. R. Facilitating Reading Comprehension through Text Structure Manipulation, June 1978.

No. 93: Ortony, A., Schallert, D. L., Reynolds, R. E., & Antos, S. J. Interpreting Metaphors and Idioms: Some Effects of Context on Comprehension, July 1978.

No. 94: Brown, A. L., Campione, J. C., & Barclay, C. R. Training Self-Checking Routines for Estimating Test Readiness: Generalization from List Learning to Prose Recall, July 1978.

No. 95: Reichman, R. Conversational Coherency, July 1978.

No. 96: Wigfield, A., & Asher, S. R. Age Differences in Children's Referential Communication Performance: An Investigation of Task Effects, July 1978.

No. 97: Steffensen, M. S., Jogdeo, C., & Anderson, R. C. A Cross-Cultural Perspective on Reading Comprehension, July 1978.

No. 98: Green, G. M. Discourse Functions of Inversion Construction, July 1978.

No. 99: Asher, S. R. Influence of Topic Interest on Black Children and White Children's Reading Comprehension, July 1978.

No. 100: Jenkins, J. R., Pany, D., & Schreck, J. Vocabulary and Reading Comprehension: Instructional Effects, August 1978.

No. 101: Shoben, E. J., Rips, L. J., & Smith, E. E. Issues in Semantic Memory: A Response to Glass and Holyoak, August 1978.

No. 102: Baker, L., & Stein, N. L. The Development of Prose Comprehension Skills, September 1978.

No. 103: Fleisher, L. S., Jenkins, J. R., & Pany, D. Effects on Poor Readers' Comprehension of Training in Rapid Decoding, September 1978.

No. 104: Anderson, T. H. Study Skills and Learning Strategies, September 1978.

No. 105: Ortony, A. Beyond Literal Similarity, October 1978.

No. 106: Durkin, D. What Classroom Observations Reveal about Reading Comprehension Instruction, October 1978.

No. 107: Adams, M. J. Models of Word Recognition, October 1978.

No. 108: Reder, L. M. Comprehension and Retention of Prose: A Literature Review, November 1978.